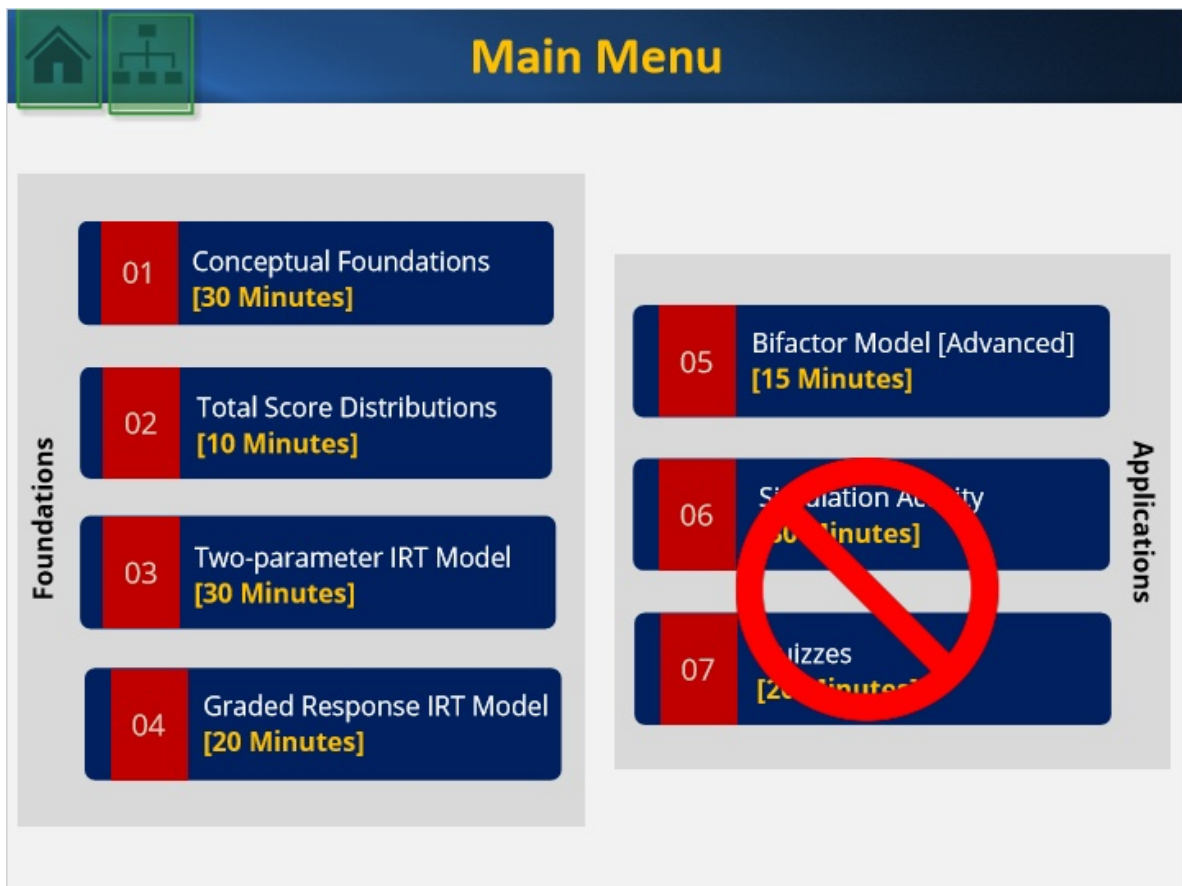


ITEMS DM13: Simulation Studies in Item Response Theory

This document contains only the core content slides from the module. In the digital module all slides can be accessed individually.

Module Organization

The module starts with an introductory section that leads to the main menu from which learners can select individual sections; this slide deck contains the slides for the five content sections only as shown below; green images in this document are clickable hotspot regions:



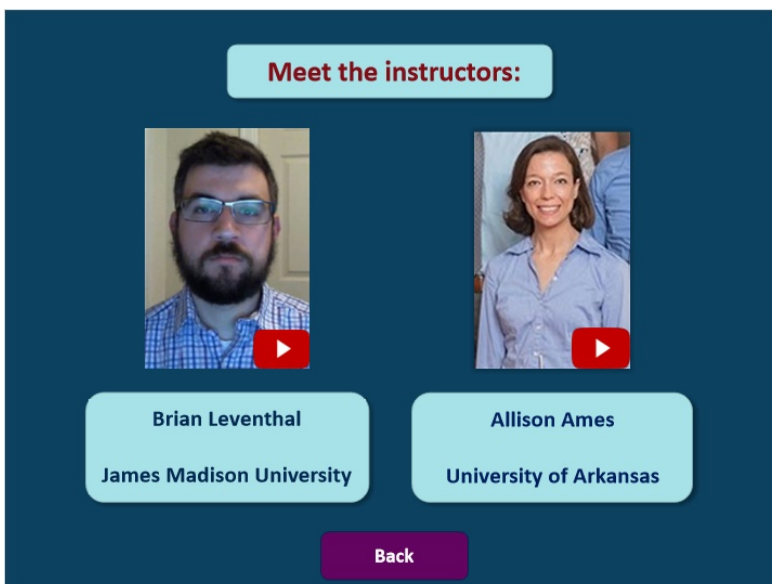
DM13 VIDEO (Section 5, Version 1.0)

1. Module Overview

1.1 Module Cover (START)


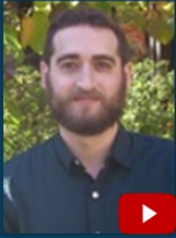


1.2 Content Team



1.3 Design Team

Meet the designers:



Jonathan Lehrfeld
ETS

André A. Rupp
Consultant

Back

1.4 Welcome



Welcome to the
ITEMS Module!

The woman to the left is Laura!

Along with the instructors
she will be guiding you
through the module content

Untitled Layer 1 (Slide Layer)

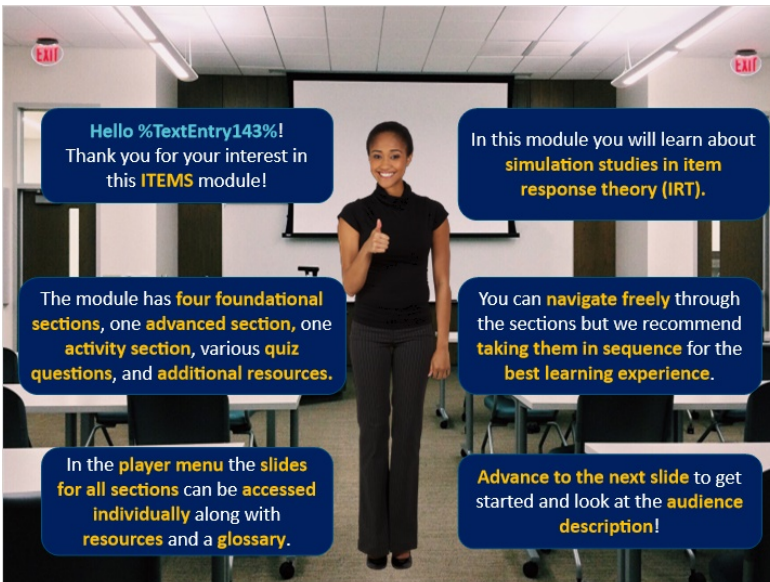


**Welcome to the
ITEMS Module!**

The woman to the left is Laura!

Along with the instructors
she will be guiding you
through the module content

1.5 Overview



Hello %TextEntry143%!
Thank you for your interest in
this ITEMS module!

In this module you will learn about
**simulation studies in item
response theory (IRT).**

The module has **four foundational
sections**, one **advanced section**, one
activity section, various **quiz
questions**, and **additional resources.**

You can **navigate freely** through
the sections but we recommend
taking them in sequence for the
best learning experience.

In the **player menu** the **slides
for all sections** can be **accessed
individually** along with
resources and a **glossary.**

Advance to the next slide to get
started and look at the **audience
description!**

1.6 Target Audience

Target Audience

Anyone who would like a gentle statistical introduction to this topic:

- graduate students and faculty in Master's, Ph.D., or certificate programs
- psychometricians and other measurement professionals
- data scientists / analysts
- research assistants or research scientists
- technical project directors
- assessment developers



However, we hope that you find the information in this module useful no matter what your official title or role in an organization is!

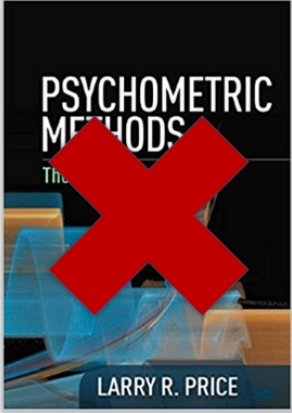
1.7 Expectations (I)



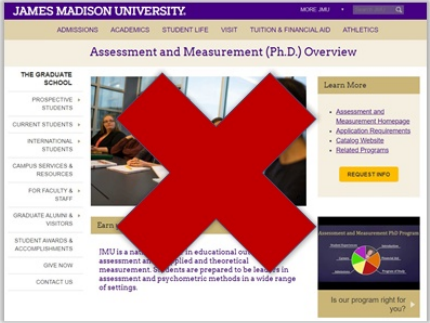
Let's discuss expectations....

1.8 Expectations (II)

ITEMS Modules in Context




PSYCHOMETRIC METHODS
The
LARRY R. PRICE



JAMES MADISON UNIVERSITY
ADMISSIONS ACADEMICS STUDENT LIFE VISIT TUITION & FINANCIAL AID ATHLETICS
Assessment and Measurement (Ph.D.) Overview
THE GRADUATE SCHOOL
PROSPECTIVE STUDENTS
CURRENT STUDENTS
INTERNATIONAL STUDENTS
CAMPUS SERVICES & RESOURCES
FOR FACULTY & STAFF
GRADUATE ALUMNI & VISITORS
STUDENT AWARDS & ACCOMPLISHMENTS
GIVE NOW
CONTACT US
Learn More
• Assessment and Measurement Homepage
• Application Requirements
• Catalog Website
• Recent Programs
REQUEST INFO
Earn
JMU is a leader in educational assessment and measurement. Students are prepared to be leaders in assessment and psychometric methods in a wide range of settings.
Assessment and Measurement PhD Program
Is our program right for you?

1.9 Learning Objectives

Learning Objectives






P₃ L₁ A₁ N₁

1. Identify the major considerations for a Monte Carlo simulation study
2. Learn important SAS procedures and techniques for data simulation
3. Adapt basic simulation techniques to IRT-specific examples
4. Apply principles from examples to more complex models and scenarios

1.10 Software Note

Software Note

The instructors use the **commercial suite SAS** for all parts of the module



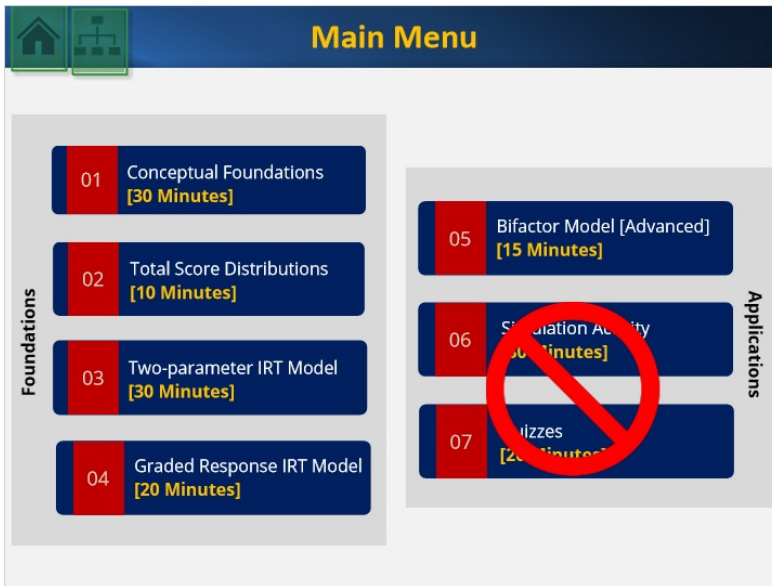
If you do not own SAS and work in **freeware suites like R or Python** you can still **learn about the principles** and **adopt code relatively easily**

1.11 Prerequisites

Prerequisites

- **Working knowledge of foundational measurement concepts:**
 - ✓ Unidimensional IRT models
 - ✓ Person parameters / latent trait parameters
 - ✓ Item parameters / thresholds
 - ✓ Response probabilities and observed scores
 - ✓ Visualizations of key relationships
- **Working knowledge of foundational simulation concepts:**
 - ✓ Bias, standard error, and mean squared error
 - ✓ Replications
 - ✓ Estimation, parameter recovery, and model fit

1.13 Main Menu



1.14 Module Cover (END)



2. Section 1: Conceptual Foundations



2.1 Module Cover (START)




2.2 Cover: Section 1



2.3 Objectives



Learning Objectives



1. Describe the primary steps and concepts of a Monte Carlo simulation study
2. Discuss each Monte Carlo simulation step in the context of item response theory
3. Identify the various simulation steps in published literature
4. Articulate what a simulation study can help a researcher accomplish - and what it cannot

2.4 Topic Selection



Simulation Foundations

Simulation Steps

Section End


2.5 Bookmark: Foundations



2.6 Important Resources

Important Resources

- General resources:
 - ➔ [Feinberg, R.R., & Rubright, J.D. \(2016\). Conducting simulation studies in psychometrics. *Educational Measurement: Issues and Practice*, 35\(2\), 36–49.](#)
 - ➔ [Harwell, M., Stone, C.A., Hsu, T., & Kirisci, L. \(1996\). Monte Carlo studies in item response theory. *Applied Psychological Measurement*, 20\(2\), 101–125.](#)
- Application of a simulation:
 - ➔ [Drasgow, F. \(1989\). An evaluation of marginal maximum likelihood estimation for the two-parameter logistic model. *Applied Psychological Measurement*, 13, 77-90.](#)

Educational Measurement 

Educational Measurement: Issues and Practice
Summer 2016, Vol. 35, No. 2, pp. 36–49



Conducting Simulation Studies in Psychometrics

Richard A. Feinberg and Jonathan D. Rubright, National Board of Medical Examiners

Simulation studies are fundamental to psychometric discourse and play a crucial role in operational and academic research. Yet, resources for psychometricians interested in conducting simulations are scarce. This Instructional Practice in Educational Measurement Series (IPEDS) module is meant to address this deficiency by providing a comprehensive introduction to the topic of simulation that can be easily understood by measurement specialists at all levels of training and experience. Specifically, this module describes the vocabulary used in simulation, reviews their applications in recent literature, and recommends specific guidelines for designing simulation studies and processing results. Additionally, an example (including computer code in R) is given to demonstrate how common aspects of simulation studies can be implemented in practice and to provide a template to help users build their own simulation.

Keywords: psychometrics, research design, simulation study

2.7 Standards



Standards for Monte Carlo Studies

Six standards:

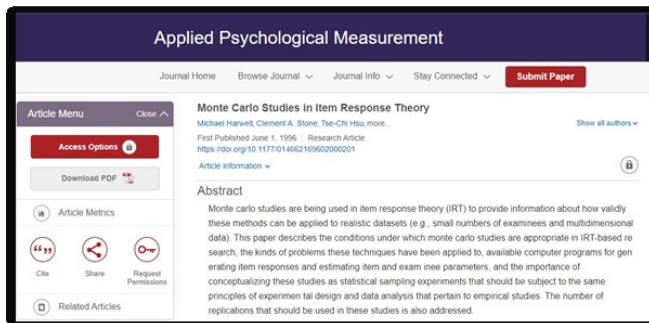
1. Can the problem could be solved analytically?
2. Is the study a minor extension of existing results?
3. Is an appropriate experimental design and analysis of MC results used?
4. Are locally-written software or modifications of public software properly documented?
5. Do the results depend on the starting values for iterative parameter estimation methods?
6. Are the choices of distributional assumptions and independent variables and their values realistic?

[Reference](#)

Reference (Slide Layer)





Reference



The screenshot shows a journal article page from 'Applied Psychological Measurement'. The article title is 'Monte Carlo Studies in Item Response Theory' by Michael Hansell, Clement A. Stone, and Tse-Chi Hsu. It was first published on June 1, 1996. The abstract discusses the use of Monte Carlo studies in Item Response Theory (IRT) to provide information about how validly these methods can be applied to realistic datasets, such as small numbers of examinees and multidimensional data. The abstract also mentions the conditions under which Monte Carlo studies are appropriate in IRT-based research, the kinds of problems these techniques have been applied to, available computer programs for generating item responses and estimating item and examinee parameters, and the importance of conceptualizing these studies as statistical sampling experiments that should be subject to the same principles of experimental design and data analysis that pertain to empirical studies. The number of replications that should be used in these studies is also addressed.

[Back](#)

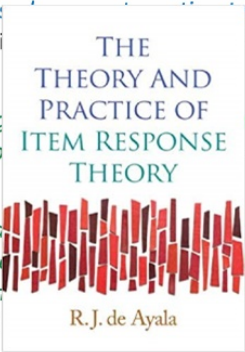
2.8 Introduction (I)

  **Introduction to Simulation Studies**



❖ “As few as 200 persons and five items were required for ‘essentially’ unbiased parameter estimates [for the 2PL].” (Drasgow, 1989; discussed in de Ayala, 2009)

❖ “The 2PL equating of local item independence must to the violation (Drasgow, 2014)

❖ “A 2:1 or larger ratio of items to parameters produced stable item parameter estimates [for the partial credit model];” (Drasgow, 1990; discussed in de Ayala, 2009)





2.9 Introduction (II)

  **Introduction to Simulation Studies**

- **How were these sample size heuristics and claims determined?**
 - ✓ Monte Carlo simulation (MCS) techniques were used.
 - ✓ In MCS, data are created by researchers based on a model and used to answer a methodological research question.
- **Applications of MCS techniques in IRT:**
 - ✓ Evaluating estimation procedures or parameter recovery,
 - ✓ Evaluating the statistical properties of an IRT-based statistic, or
 - ✓ Comparing methodologies used in conjunction with IRT.

2.10 Example: Bias

**Example: Estimation Bias**



Evaluating estimation procedures or parameter recovery

Research Question:

Are the item parameter estimates derived from this new estimation procedure, on average, close to their true values at varying sample sizes?


Illustration:
Bias vs. Unbiased

Introduction to Simulation Studies (Slide Layer)


**Introduction to Simulation Studies**

If the center of the target is the “true” population parameter. When estimates are clustered around the “true” parameter, the estimator is said to be unbiased.

Biased





Unbiased



Back

2.11 Example: Type-I Error

  **Example: Properties of a Statistic**

Evaluating the statistical properties of an IRT-based statistic
For example, a goodness-of-fit statistic is evaluated

Research Question:
What is the type-I error rate of this statistic?

Illustration:
Type-I Error

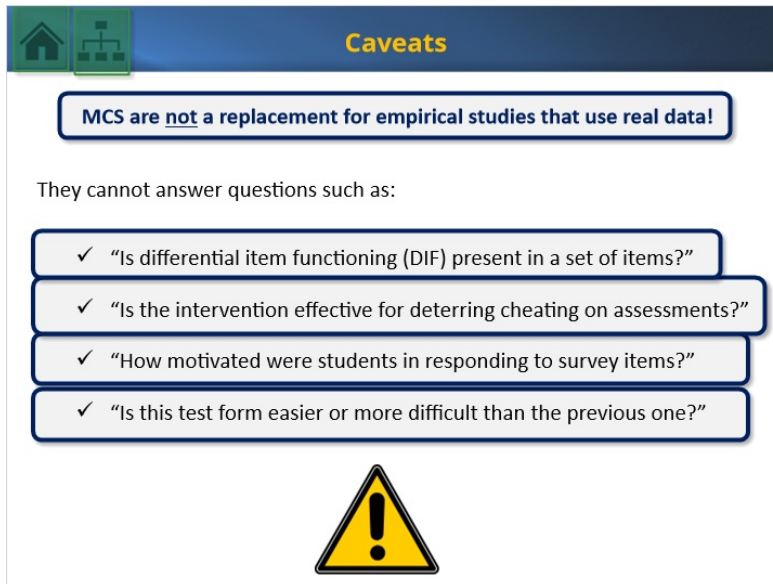
Type-I Error (Slide Layer)

  **Type-I Error**

	Finding: No Misfit	Finding: Misfit
True Model: No Misfit	Correct Decision! ✓	Type I Error! ⚠
True Model: Misfit	Type II Error! ⚠	Correct Decision! ✓ (Power)

Back

2.12 Caveats




The slide features a dark blue header with a home icon and a tree icon on the left, and the word "Caveats" in yellow on the right. Below the header, a white box with a blue border contains the text "MCS are not a replacement for empirical studies that use real data!". Underneath, it says "They cannot answer questions such as:" followed by four white boxes with blue borders, each containing a checkmark and a question. At the bottom center is a yellow warning triangle with a black exclamation mark.

MCS are not a replacement for empirical studies that use real data!

They cannot answer questions such as:

- ✓ "Is differential item functioning (DIF) present in a set of items?"
- ✓ "Is the intervention effective for deterring cheating on assessments?"
- ✓ "How motivated were students in responding to survey items?"
- ✓ "Is this test form easier or more difficult than the previous one?"



2.13 Bookend: Foundations



The slide has a dark blue background. In the top left corner, there is a blue box with a white home icon and a white tree icon. On the left side, a woman in a black top and pants is giving a thumbs-up. To her right, the text "This is the end of this part." is displayed. Below the text is a purple rounded rectangle with the words "Topic Selection" in white.

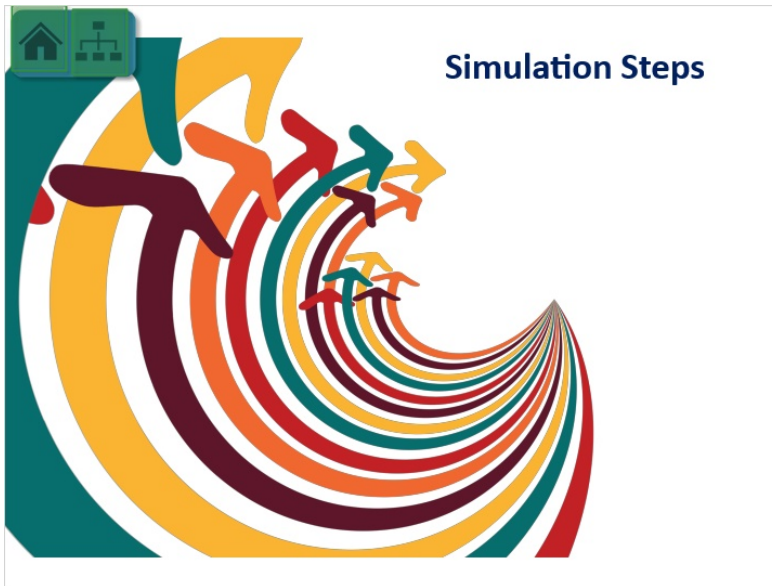
 



This is the end of this part.

Topic Selection

2.14 Bookmark: Simulation Steps



2.15 Step Selection

An "Overview" slide showing a list of eight simulation steps. A large red arrow on the left points downwards. A blue book icon is next to step 2, and a person at a computer icon is next to step 5. A "Topic Selection" button is in the bottom right, and a "Click on each row to learn more" button is at the bottom center.



Overview

1. Specifying the IRT research question(s)
2. Defining and justifying conditions
3. Specifying the experimental design and outcome(s) of interest
4. Simulating data under the specified conditions
5. Estimating parameters
6. Comparing true and estimated parameters
7. Replicating the procedure a specified number of times
8. Analyzing results based on the design and research questions

Click on each row to learn more

Topic Selection


2.16 Research Questions (I)

**Step 1: Research Question(s)**



The research question will guide all other aspects of the simulation

➔ *“Articulating a clear research question ... forces the researcher to ensure that the exact design choices made align with the question being asked.”*
Feinberg & Rubright, 2016

➔ *“... relies heavily on knowledge of a literature.”*
Harwell, Stone, Hsu, & Kirisci, 1996



2.17 Research Questions (II)

**Step 1: Example**

Drasgow (1989) provides background from a literature review:

- Large sample sizes and long tests
- Common situations in which these are unrealistic or untenable

Example research question from Drasgow (1989):

What is the range of conditions that allow accurate calibration of two-parameter logistic items by marginal maximum likelihood estimation?

[Reference](#)

Reference (Slide Layer)



An evaluation of marginal maximum likelihood estimation for the two-parameter logistic model.

EXPORT Add To My List Full text from publisher

Database: APA PsycInfo Journal Article

[Cite](#) [Full text from publisher](#)

Citation

Dragow, F. (1989). An evaluation of marginal maximum likelihood estimation for the two-parameter logistic model. *Applied Psychological Measurement*, 13(1), 77-90. <https://doi.org/10.1177/014662168901300108>

Abstract


Investigated the accuracy of marginal maximum likelihood estimates of the item parameters of the 2-parameter logistic model. Estimates were obtained for 4 sample sizes (200, 300, 500, and 1,000) and 4 test lengths (5-, 10-, 15-, and 25-items); joint maximum likelihood estimates were also computed for the 2 longer test lengths. Each condition was replicated 10 times, which allowed evaluation of the accuracy of estimated item characteristic curves, item parameter estimates, and estimated standard errors of item parameter estimates for individual items. Results show that items typical of widely used job satisfaction scale and moderately easy tests had satisfactory marginal estimates for all sample sizes and test lengths. Larger samples were required for items with extreme difficulty or discrimination parameters. Marginal estimation was substantially better than joint maximum likelihood estimation. (PsycINFO Database Record (c) 2016 APA, all rights reserved)

Back

2.18 Bookend: Step 1

A dark blue slide with a woman in a black top and pants giving a thumbs up. In the top left corner are a home icon and a tree icon. The text "This is the end of this step." is on the right. A purple button with "Step Selection" is at the bottom right.

2.19 Justifying Conditions





Step 2: Justification

- The research questions dictate the independent variables and resulting conditions to be included in the simulation
- Conditions must be justified and clearly delineated

Example from Drasgow (1989):

- Test length: 5 to 25 items
- Sample size: 200 to 1,000 simulated examinees
- Justification: These ranges encompass the values seen in many applied studies at the time



2.20 Bookend: Step 2



This is the end of this step.

Step Selection

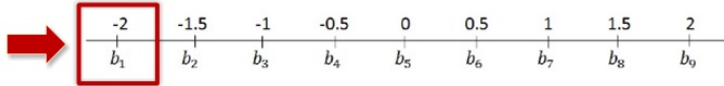
2.21 Experimental Design (I)

  **Step 3: Experimental Conditions (I)**

“Model parameters also represent an independent variable in an MCS”
Harwell, Stone, Hsu, & Kirisci, 1996



Fixed effect: Parameters are represented as (often equally-spaced) values across a fixed range or as estimates from a previously calibrated test

- **Example:** 9 b parameters are evenly spaced from -2 to 2
- **Advantage:** Simple setup with known values
- **Disadvantage:** Limited generalizability



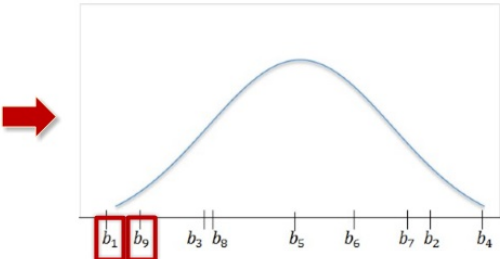
A horizontal number line ranging from -2 to 2 with tick marks at intervals of 0.5. The points are labeled b_1 through b_9 from left to right. A red arrow points to the point b_1 at -2, which is also enclosed in a red box.

2.22 Experimental Design (II)

  **Step 3: Experimental Conditions (II)**



Random effect: Parameter values are sampled from a specified distribution

- **Example:** $b \sim N(0, \sigma^2 = 4)$
- **Advantage:** Some generalizability is obtained
- **Disadvantage:** Unusual combinations of parameters might occur



A graph showing a normal distribution curve. The x-axis is labeled with points b_1 through b_9 from left to right. A red arrow points to the left side of the curve, and a red box highlights the points b_1 and b_9 .

2.23 Experimental Design (III)



  **Step 3: Experimental Conditions (III)**

“Researchers also must consider the relationship between the number of independent variables, the efficiency of the study, and the interpretability of the results.”
Harwell, Stone, Hsu, & Kirisci, 1996

Example from Drasgow (1989)

- Test length (4 levels: 5, 10, 15, 25)
- Sample size (4 levels: 200, 300, 500, 1000)
- Item parameters (2 levels: typical, extreme)
- 4 x 4 x 2 fully crossed design = 32 total combinations

2.24 Experimental Design (IV)



  **Step 3: Experimental Conditions (IV)**

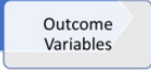

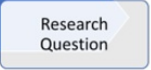
Type and number of independent variables guide selection of experimental design

Example from Drasgow (1989)


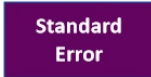

- **Factorial design:** Small number of independent variables with relatively few values
- **Investigative focus:** Each combination of the conditions is important and will be used in the simulation




2.25 Experimental Design (V)

Step 3: Outcome Measures





Example: Accuracy of parameter estimation
Outcome: How well the “true” parameter value is recovered





Bias (Slide Layer)

Outcome: Bias

- Average deviance between the “true” and estimated values of the parameter
- Systematic error

$$\frac{\sum_{i=1}^n (\hat{\theta}_i - \theta_{True})}{n}$$

$\hat{\theta}_i$: Parameter estimate from replication i
 θ_{True} : True parameter value
 n : Total number of replications

Back

Standard Error (Slide Layer)



Outcome: Standard Error

- Standard deviation across estimated values
- Random error

$$\sqrt{\frac{\sum_{i=1}^n (\hat{\theta}_i - \bar{\theta})^2}{n - 1}}$$

$\bar{\theta}$ is the average parameter value

$\hat{\theta}_i$ is the estimated parameter value of replication i

n is the replication

Back

Mean Squared Error (Slide Layer)



Outcome: Mean Square Error

- Average squared deviance between the estimated and true values
- Total error

$$\sqrt{\frac{\sum_{i=1}^n (\hat{\theta}_i - \theta_{True})^2}{n - 1}}$$

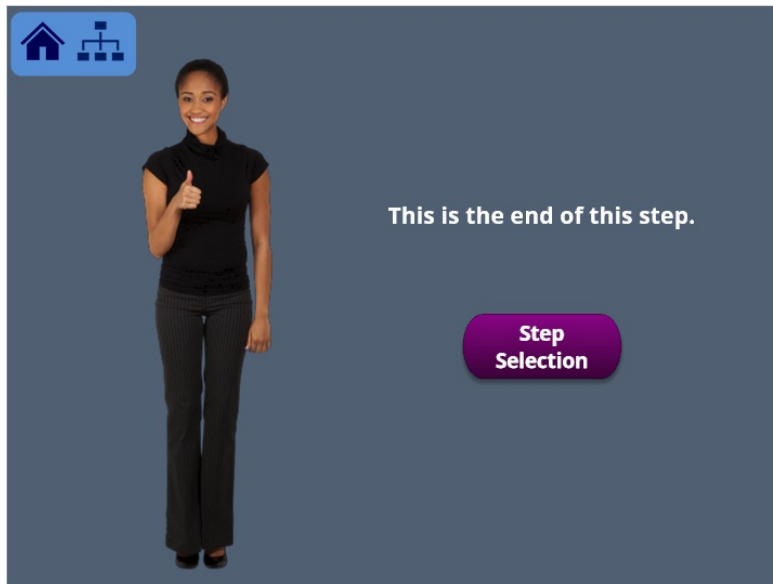
$\hat{\theta}_i$: Parameter estimate from replication i

θ_{True} : True parameter value

n : Total number of replications

Back

2.26 Bookend: Step 3

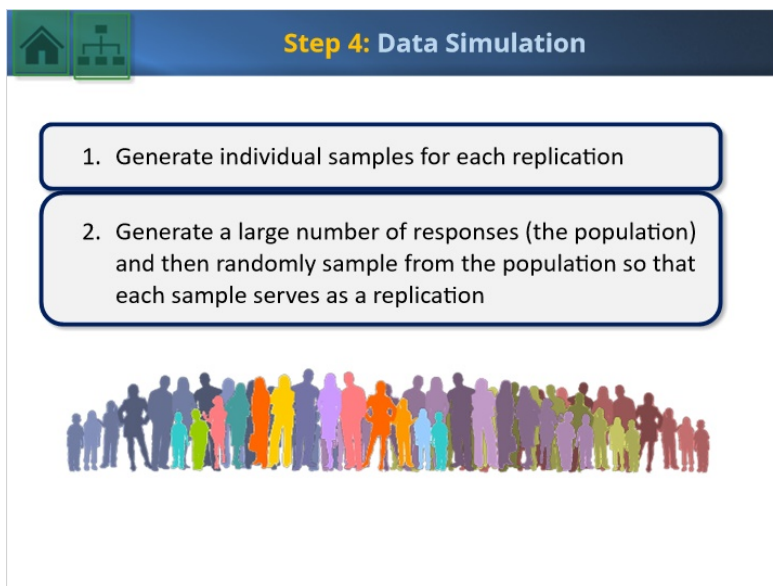


This is the end of this step.

Step Selection


The image shows a woman in a black top and pants giving a thumbs up. To her right is a purple button with the text 'Step Selection'. Above the woman is a blue icon containing a house and a tree diagram. The background is a dark blue gradient.

2.27 Simulating Data (I)





Step 4: Data Simulation

1. Generate individual samples for each replication
2. Generate a large number of responses (the population) and then randomly sample from the population so that each sample serves as a replication




The slide features a dark blue header with a green icon of a house and tree, and the text 'Step 4: Data Simulation'. Below the header are two rounded rectangular boxes containing the numbered steps. At the bottom is a colorful silhouette of a diverse group of people.

2.28 Simulating Data (II)



  **Step 4: Person Parameters**

Simulate each individual's latent ability

Common Choice: $\theta \sim N(0,1)$

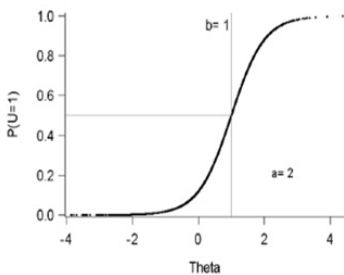


2.29 Simulating Data (III)

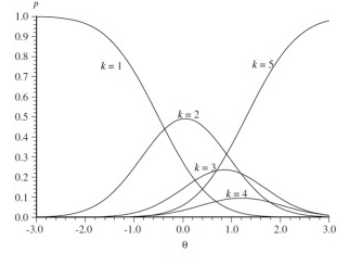
  **Step 4: Response Probabilities**

Combine latent ability with item parameters to obtain a matrix of response probabilities



Dichotomous IRT



Polytomous IRT

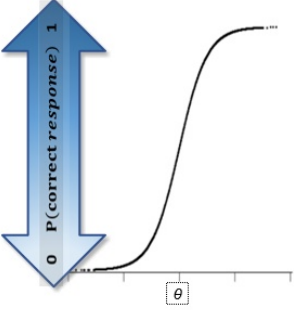


2.30 Simulating Data (IV)



  **Step 4: Dichotomous Responses**


Transform probabilities into responses

- $P(\text{correct response}) = \text{high}$
Response = 1
Correct
- $P(\text{correct response}) = \text{low}$
Response = 0
Incorrect



2.31 Bookend: Step 4



 



This is the end of this step.

Step Selection



2.32 Estimating Parameters (I)



Step 5: Parameter Estimation

- **Available tools:**
 - ✓ Commercially available programs (e.g., Winsteps or Bilog)
 - ✓ Open source packages (e.g., R package ltm or Stan)
 - ✓ Their own routine (e.g., Fortran, SAS, or Basic)
- **Adequacy** of the estimation algorithm must be documented
- **Validation evidence examples:**
 - ✓ Using the program to analyze a well-known dataset
 - ✓ Analyzing item responses with perfect fit to an item response function in which case the estimation routine should return the exact item parameters

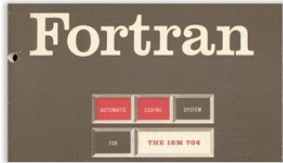
2.33 Estimating Parameters (II)




Step 5: Example

Example from Drasgow (1989)

- A FORTRAN program was written to estimate the a (discrimination) and b (difficulty) parameters of the 2PL
- Validity evidence was not provided



2.34 Estimating Parameters (III)





Step 5: Addendum

Example from Drasgow (1989)

Interest was in parameter recovery for the 2PL so data were generated from the 2PL

In some MCS studies the data-generating model is more complex than the analysis model so that misfit is introduced (e.g., Orlando & Thissen, 2000)


2.35 Bookend: Step 5



This is the end of this step.

Step Selection

2.36 Outcome Measures





Step 6: Outcome Measures

Example from Drasgow (1989)

- Bias of discrimination and difficulty parameter estimates
- Distance between the true 2PL response curve and the estimated response curve
- Comparison of observed and mean estimated standard errors



2.37 Bookend: Step 6



This is the end of this step.

Step Selection

2.38 Replications






Step 7: Replications

“More replications are always better in terms of producing a more accurate and reliable estimate of the parameters of interest.” Feinberg & Rubright, 2016

The purpose of the study has an important effect on the number of replications selected.

Study Purpose	Replications
Parameter Recovery	500-1000+
Bayesian MCS	10-100
Comparing IRT-based methodologies	10-100



2.39 Bookend: Step 7



This is the end of this step.

Step Selection


2.40 Analyzing Results

**Step 8: Analysis of Results**



The independent and outcomes variables will dictate the analysis

Example from Drasgow (1989)

- No inferential results were reported
- Summary statistics were graphically displayed



2.41 Bookend: Step 8







This is the end of this step.

Step Selection



2.42 Simulation Standards



Standards for Monte Carlo Studies

1. Can the problem could be solved analytically?
2. Is the study a minor extension of existing results?
3. Is an appropriate experimental design and analysis of MC results used?
4. Are locally-written software or modifications of public software properly documented?
5. Do the results depend on the starting values for iterative parameter estimation methods?
6. Are the choices of distributional assumptions and independent variables and their values realistic?

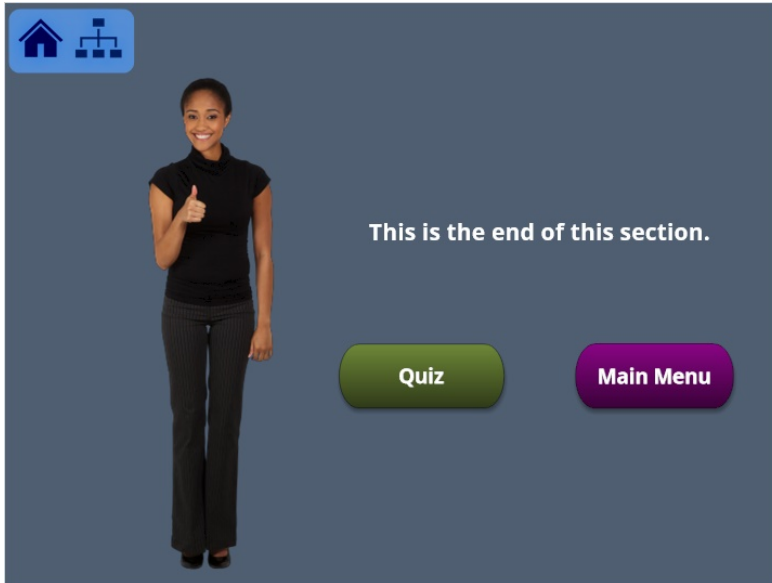
2.43 Summary



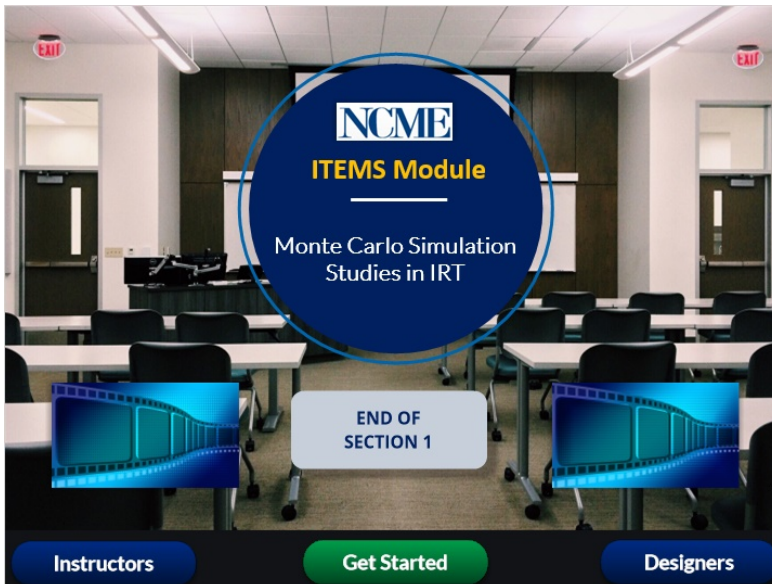
Summary

- Important steps of Monte Carlo simulation studies:
 - ✓ Always begins with a research question, and all aspects follow from the question
 - ✓ Simulations are not a substitution for empirical work as they are designed to answer methodological questions
- Demonstrated all steps from Drasgow's 1989 parameter recovery study
- Next steps: implement the software steps in SAS

2.44 Bookend: Section 1



2.45 Module Cover (END)

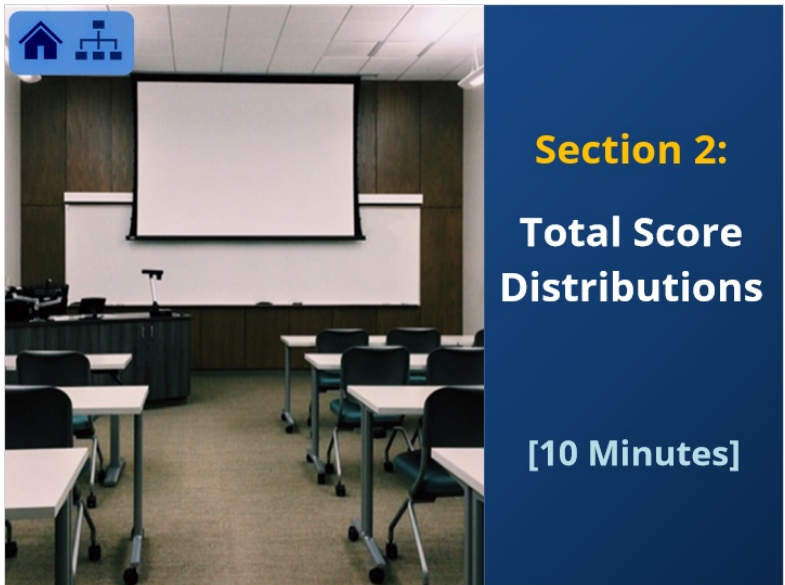


3. Section 2: Total Score Distributions



3.1 Module Cover (START)



3.2 Cover: Section 2



3.3 Objectives

  **Learning Objectives**





1. Identify situations for total score simulations

2. Articulate rationale behind total score simulation methods

3. Describe the process for total score data simulation

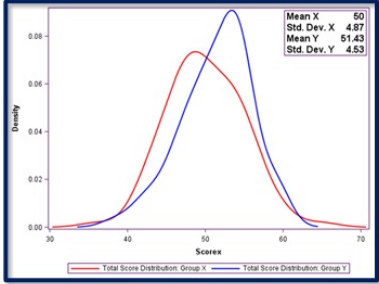
4. Modify example code for specific research scenarios

3.4 Total Score Distributions

  **Total Score Distributions**

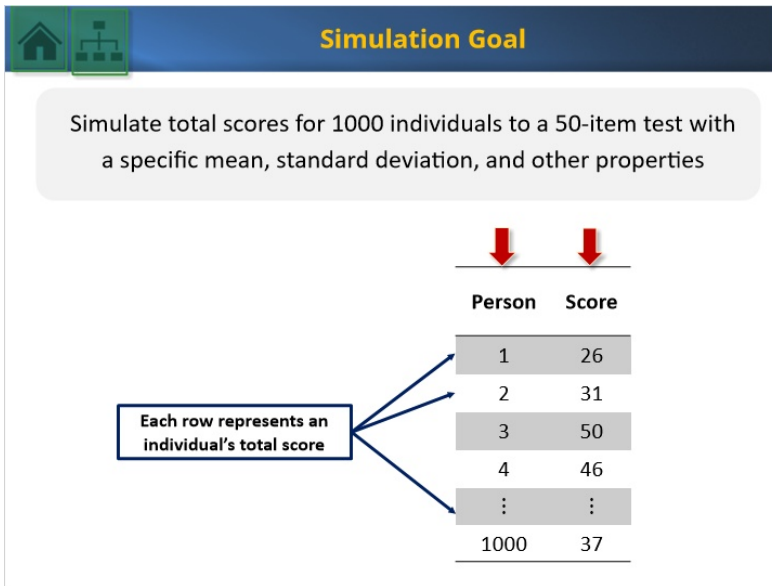
Total score:
sum of all the scored items

Total score distribution:
distribution of the sample's total scores



Statistic	Group X	Group Y
Mean	50	51.43
Std. Dev.	4.87	4.63

3.5 Simulation Goal



Simulation Goal

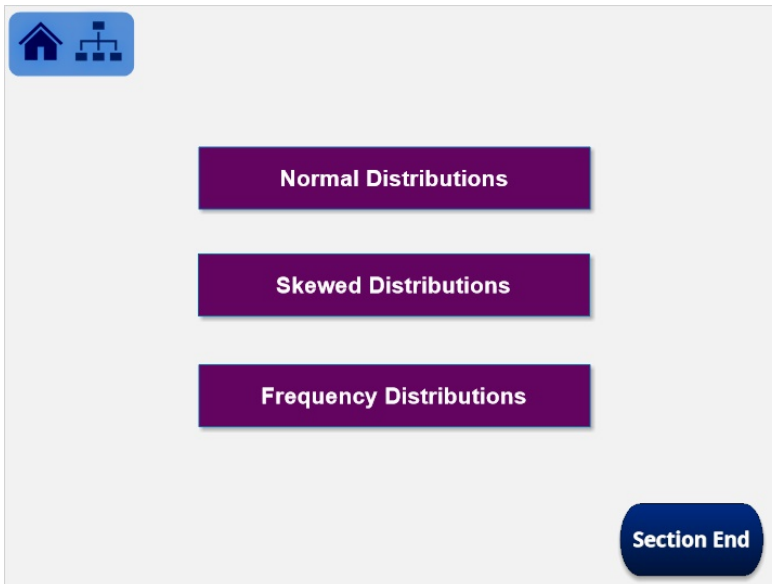
Simulate total scores for 1000 individuals to a 50-item test with a specific mean, standard deviation, and other properties

Person	Score
1	26
2	31
3	50
4	46
⋮	⋮
1000	37

Each row represents an individual's total score

The interface features a dark blue header with a home icon and a tree icon. Below the header is a light gray box containing the simulation goal text. A table with two columns, 'Person' and 'Score', is shown with red arrows pointing to the column headers. A callout box with a blue border and white background points to the first four rows of the table, stating 'Each row represents an individual's total score'. The table rows are: (1, 26), (2, 31), (3, 50), (4, 46), (⋮, ⋮), and (1000, 37).

3.6 Topic Selection



Normal Distributions

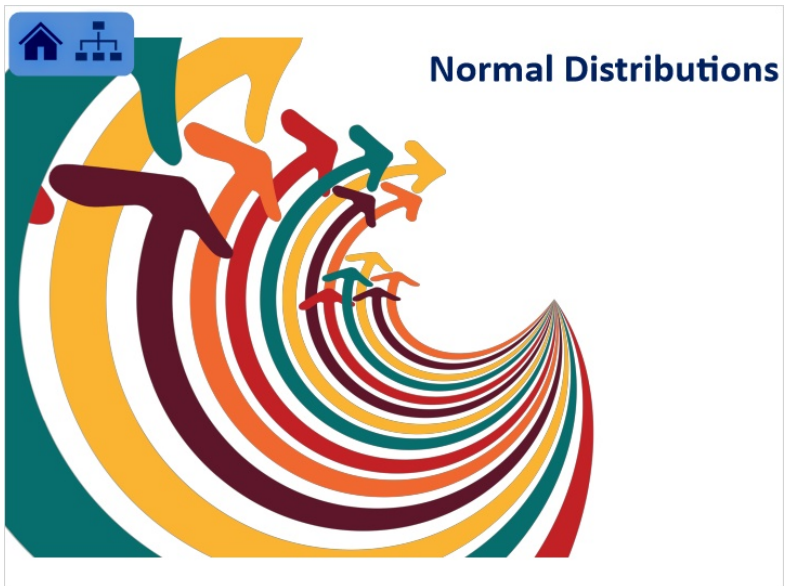
Skewed Distributions

Frequency Distributions

Section End

The interface features a light gray background with a blue header containing a home icon and a tree icon. Three purple rectangular buttons are stacked vertically, labeled 'Normal Distributions', 'Skewed Distributions', and 'Frequency Distributions'. A dark blue rounded rectangular button labeled 'Section End' is located in the bottom right corner.

3.7 Bookmark: Normally Distributed Data



3.8 Standard Normal (I)

Generate Random Normal Values

```
x = rand("Normal", 50, 5);
```

Yields normally distributed data (mean = 50, standard deviation of = 5)

Result:

person	x
1	56.2034
2	44.9303
3	48.3771
4	46.6160
5	51.9494
6	53.7597
...	...

Test scores are typically recorded in whole numbers, not decimals.

We want:

person	x
1	56
2	45
3	48
4	47
5	52
6	54
...	...

3.9 Standard Normal (II)

Truncate Generated Values

```
x = rand("Normal", 50, 5);  
Scorex=round(x, 1);
```


round() : Rounds the variable or constant to the nearest integer.

person	x	Scorex
1	56.2034	56
2	44.9303	45
3	48.3771	48
4	46.6160	47
5	51.9494	52
6	53.7597	54
...

3.10 Bookend: Normal Distributions

This is the end of the this part.

Topic Selection



3.11 Bookmark: Skewed Data



3.12 Skewed Distribution (I)

Skewness Formula

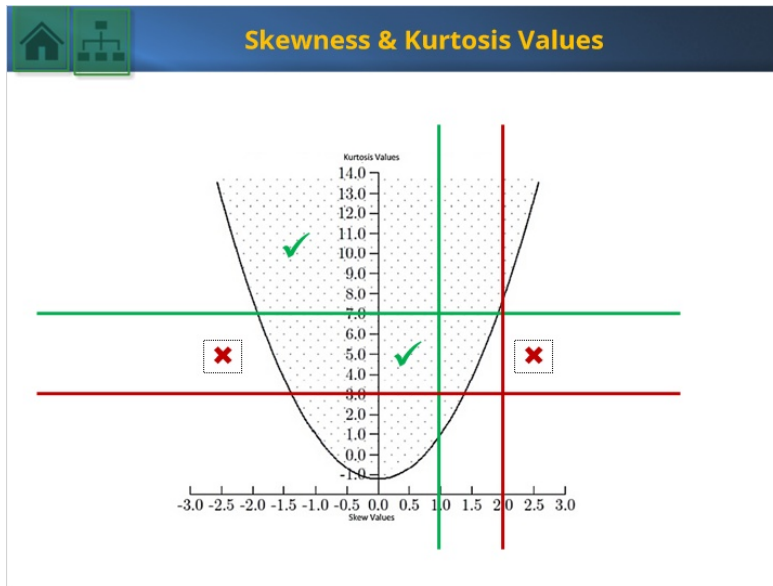
- Generate a random **standard normal value X** and compute:
$$skewX = -c + bX + cX^2 + dX^3$$
- b , c , and d come from **Fleishman's power method** implemented via coefficient table or *Proc NLP*

Example

skew = 1, kurtosis = 5
 $b = 0.108968$
 $c = 0.730401$
 $d = 0.079435$

Two graphs illustrating skewed distributions. The left graph is labeled "Negative Skew" and shows a distribution with a long tail on the left side. The right graph is labeled "Positive Skew" and shows a distribution with a long tail on the right side. Red arrows point to the peaks of both distributions.

3.13 Skewed Distribution (II)



3.14 Skewed Distribution (III)

The slide is titled "Skewed Distribution". It contains the following SAS code:

```
%let n_person=100;

data TotalScore;
do person = 1 to &n_person;
x = rand("Normal",0,1);
skewx=-.108968+.730401*x+.108968*x**2+.079435*x**3;
output;
end;
run;
```

Below the code is a purple button labeled "Results".

Results (Slide Layer)

  **Adding Skew**

```
proc means data=TotalScore skew kurt mean std;  
var skewx;  
run;
```

Skewness	Kurtosis	Mean	Std. Dev.
1.021336	5.084481	0.005539	1.008835

These may not be the mean and standard deviation that we want so transformations are necessary

Back

3.15 Skewed Distribution (IV)

  **Transformation & Rounding**

```
%let n_person=100;  
%let sd=5;  
%let mean=50;  
  
data TotalScore;  
do person = 1 to &n_person;  
x = rand("Normal",0,1);  
skewx=-.108968+.730401*x+.108968*x**2+.079435*x**3;  
score=round(skewx*&sd+&mean,1);  
output;  
end;  
run;
```

$$X = Z * \sigma + \mu$$

Results

Results (Slide Layer)

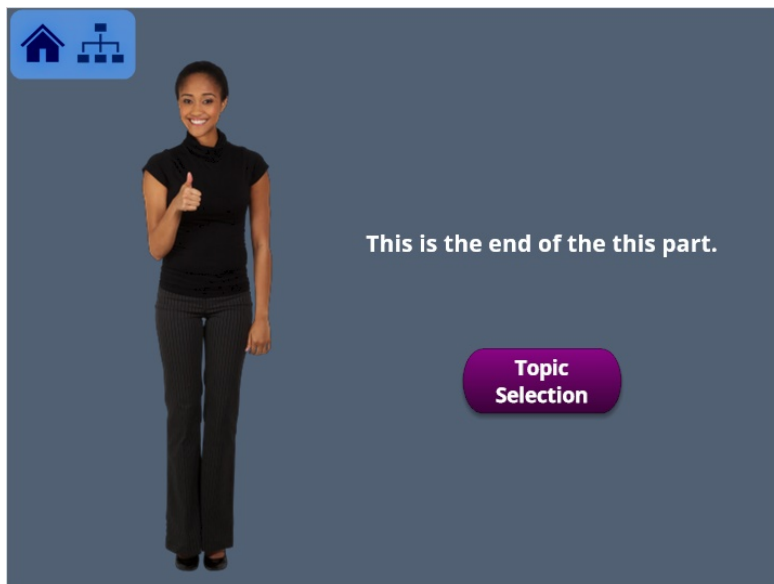


```
proc means data=TotalScore skew kurt mean std;  
var skewx score;  
run;
```

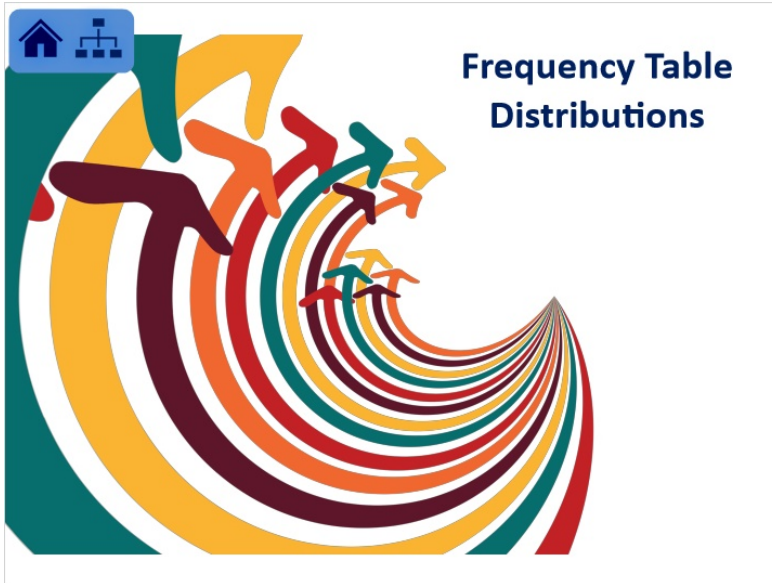
Variable	Skewness	Kurtosis	Mean	Std. Dev.
skewx	1.021336	5.084481	0.005539	1.008835
score	1.014059	5.009353	50.0309	5.049544

Back

3.16 Bookend: Skewed Distributions



3.17 Bookmark: Data from Freq Distribution



3.18 Frequency Table (I)

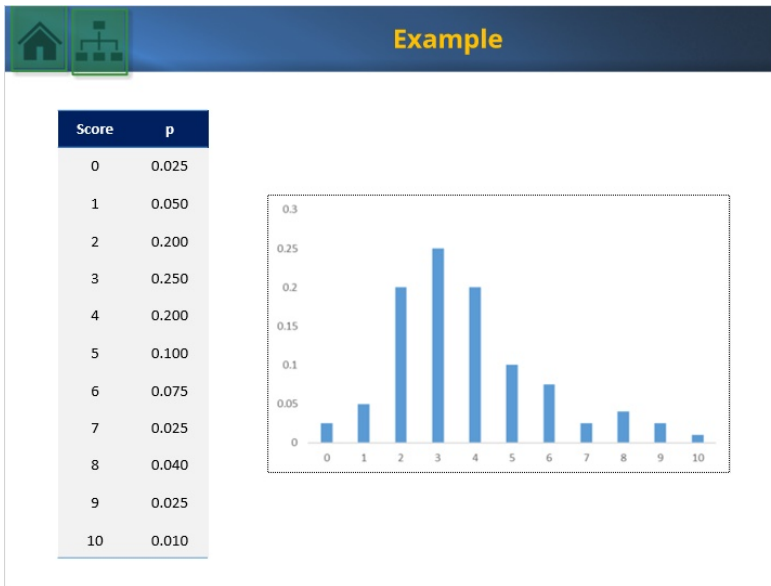
Generating Values from a Frequency Table

- Simulate discrete values from a frequency distribution
- The tabled distribution takes on the values $1, 2, \dots, n$ with specified probabilities

```
rand("Table", p1, p2...);
```

A small icon of a table with a grid pattern.

3.19 Frequency Table (II)



3.20 Frequency Table (III)

Implementation

```
data CatScore;
do person = 1 to &n_person;
Score = rand("Table",
0.025,0.05,0.2,0.25,0.2,0.1,0.075,0.025,0.04,0.025,0.01);
Scorex=Score-1;
output;
end;
run;
```

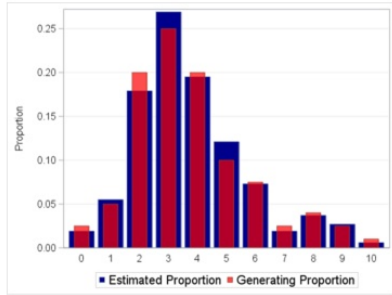
Vector of the probability of scores from 0 to 1

SAS simulates scores from 1 to 11 but we want them from 0 to 10 so we need to subtract 1

Results

Results (Slide Layer)

Generating Values from a Frequency Table



Score x	Estimated proportion	Generating proportion
0	0.019	0.025
1	0.055	0.050
2	0.179	0.200
3	0.269	0.250
4	0.195	0.200
5	0.121	0.100
6	0.073	0.075
7	0.019	0.025
8	0.037	0.040
9	0.027	0.025
10	0.006	0.010

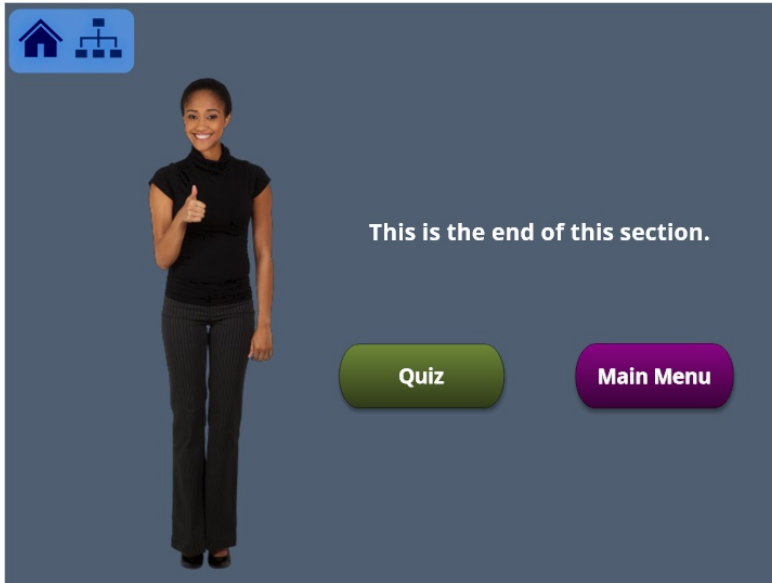
Back

3.21 Bookend: Frequency Table Distributions

This is the end of the this part.

Topic Selection

3.22 Bookend: Section 2



3.23 Module Cover (END)

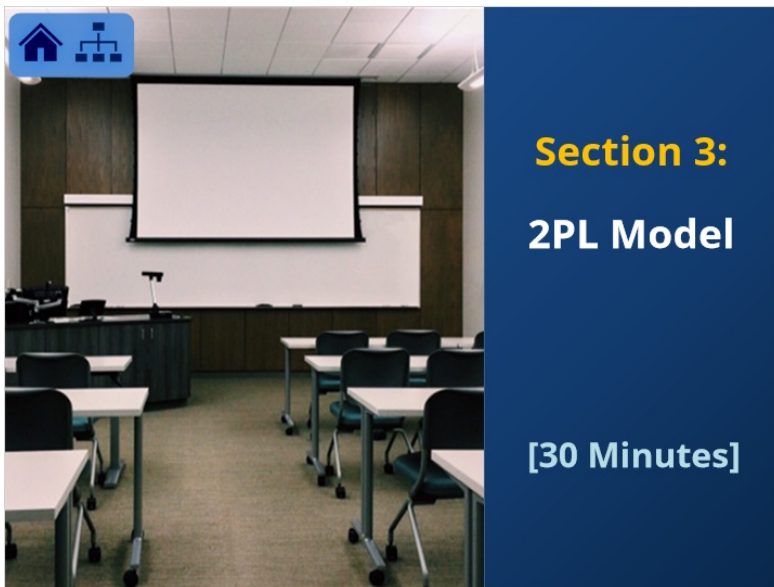


4. Section 3: Two-parameter Model



4.1 Module Cover (START)




4.2 Cover: Section 3



4.3 Objectives





Learning Objectives



1. Identify key parameters in the 2 PL model for simulation
2. Describe methods of how to chose generating parameters
3. Articulate the process for item response simulation using the 2PL
4. Modify example code for specific research scenarios

4.4 Simulation Goal



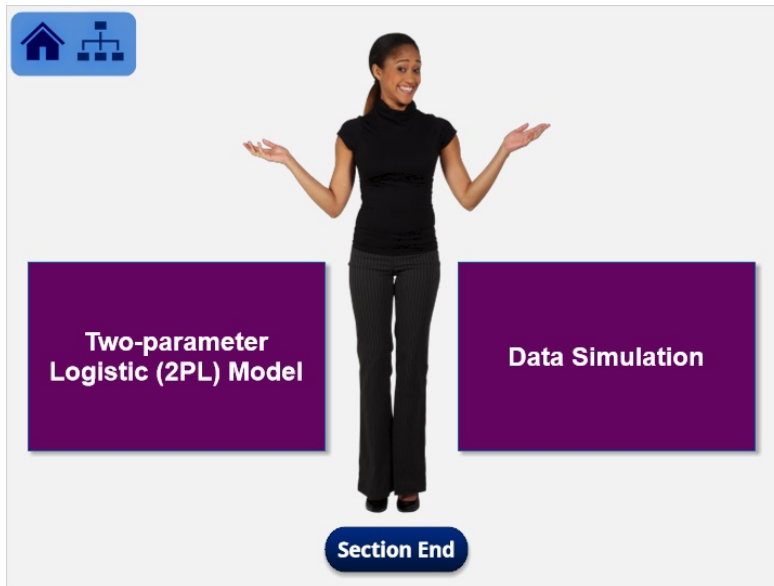
Simulation Goal

- Simulate item responses for 100 individuals to a 14-item test
- Each item is going to be dichotomously scored as
 - ✓ 1 = Correct
 - ✓ 0 = Incorrect
- Goal dataset:

	Item 1	Item 2	Item 3	...	Item 14
1	1	1	0	...	1
0	1	1	1	...	0
⋮	⋮	⋮	⋮	⋮	⋮
0	1	1	1	...	0

Each row represents an individual's response set



4.5 Topic Selection



4.6 Bookmark: Two-paramter Model





4.7 Two-parameter Model (I)

  **Two-Parameter Logistic (2PL) Model (I)**

$$P_j(U_j = 1 | \omega_j, \theta) = \frac{1}{1 + e^{-Da_j(\theta - b_j)}}$$

- Probability an individual with a given ability θ correctly answers item j that has item parameters ω_j
- U_j : response on item j (1 = correct, 0 = incorrect)
- ω_j is the vector of all item parameters
- θ is an ability value on the trait of interest

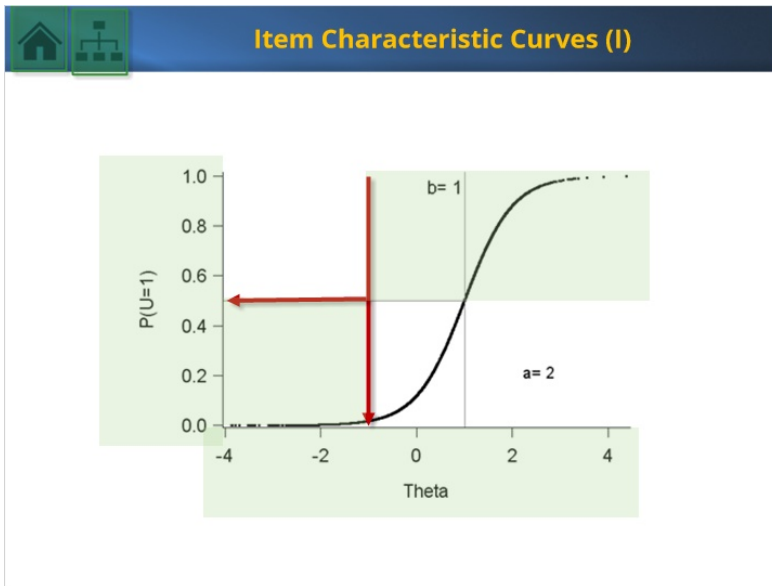
4.8 Two-parameter Model (II)

  **Two-Parameter Logistic (2PL) Model (II)**

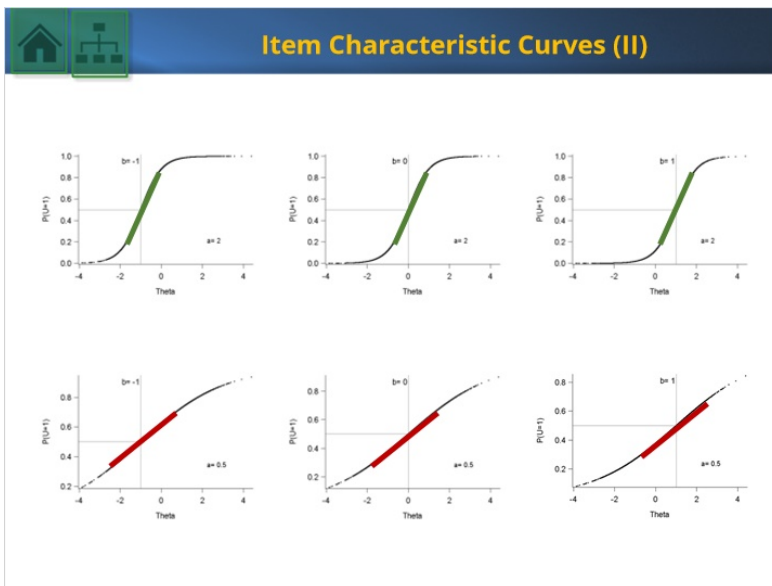
$$P_j(U_j = 1 | \omega_j, \theta) = \frac{1}{1 + e^{-Da_j(\theta - b_j)}}$$

- e is Euler's number (approx. 2.718)
- D is the scaling parameter
 - ✓ Here we set $D = 1$
 - ✓ Other common value is $D = 1.702$
- a_j is the discrimination parameter for item j
- b_j is the difficulty/threshold parameter for item j

4.9 Item Characteristic Curves (I)



4.10 Item Characteristic Curves (II)



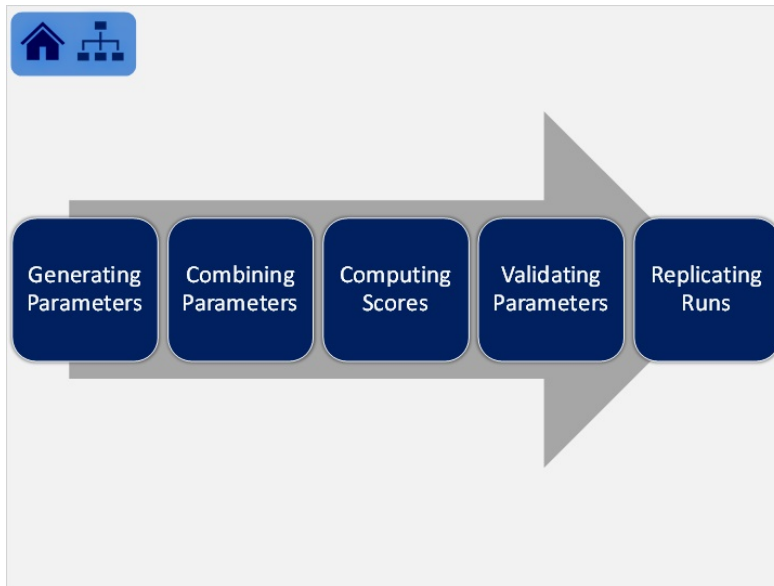
4.11 Bookend: Two-parameter Model



4.12 Bookmark: Generating Parameters with 2PLM





4.13 Step Selection



4.14 Person Parameters

```
%let n_person = 100;
data person_parm;
  do i = 1 to &n_person;
    theta=rand('normal',0,1);
    output;
  end;
  keep theta;
run;
```

4.15 Item Parameters






Item Parameters

- Create a dataset that contains item parameters.
- This may be based off of empirical estimates, previous literature, etc.
- You may randomly select item parameters as well.
- Literature can be helpful in determining what distribution to randomly select from.

```
data item_parms;  
  input a b;  
  cards;  
  1.7 -2  
  1.7 -1  
  1.7 -.5  
  1.7 0  
  1.7 .5  
  1.7 1  
  1.7 2  
  1.9 -2  
  1.9 -1  
  1.9 -.5  
  1.9 0  
  1.9 .5  
  1.9 1  
  1.9 2  
  ;  
run;
```

4.16 Bookend: Step 1



This is the end of this part.

Topic Selection

4.17 Parameter Matrix (I)

Parameter Matrix (I)

- We need in the same data set / parameter matrix:
 - ✓ an individual's ability (θ)
 - ✓ item parameters (a_j and b_j)
- We want our data set / parameter matrix to look like this:

Theta	a_1	b_1	...	a_{14}	b_{14}
-1.43	1.7	-2	...	1.9	2
2.81	1.7	-2	...	1.9	2
⋮	⋮	⋮	⋮	⋮	⋮
0.09	1.7	-2	...	1.9	2

4.18 Parameter Matrix (II)

Parameter Matrix (II)

	a	b
1	1.7	-2
2	1.7	-1
3	1.7	-0.5
4	1.7	0
5	1.7	0.5
6	1.7	1
7	1.7	2
8	1.9	-2
9	1.9	-1
10	1.9	-0.5
11	1.9	0
12	1.9	0.5
13	1.9	1
14	1.9	2

```
proc transpose data=item_parms out=a_wide prefi:
var a;
run;
```

NAME OF FORMER VARIABLE	a1	a2	a3	a4	a5	a6	a7
1 a	1.7	1.7	1.7	1.7	1.7	1.7	1.7

```
proc transpose data=item_parms out=b_wide prefi:
var b;
run;
```

NAME OF FORMER VARIABLE	b1	b2	b3	b4	b5	b6	b7	b8
1 b	-2	-1	-0.5	0	0.5	1	2	2

```
data items_wide;
merge a_wide b_wide;
drop _name_;
run;
```

Goal:

a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	b1	b2	b3
1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.7	1.9	1.9	1.9	1.9	1.9	1.9	-2	-1	-0.5

4.19 Parameter Matrix (III)

Parameter Matrix (III)

```

data all_parms;
  set person_parm;
  If _N_=1 then set items_wide ;
run;
    
```

parm	a1	a2	a3	a4	a5	a6	a7	a8	a9	a10	a11	a12	a13	a14	b1	b2	b3
1	0.37776205	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
2	0.48442804	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
3	1.388117271	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
4	0.500362046	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
5	1.388117271	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
6	1.42895431	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
7	1.163679787	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
8	0.48882147	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
9	0.48882147	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
10	0.29879462	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
11	0.02346381	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
12	1.19138548	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
13	1.388117271	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
14	0.500362046	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
15	0.89882814	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
16	0.29879462	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
17	0.48442804	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
18	0.89882814	17	17	17	17	17	17	17	19	19	19	19	19	19	2	-1	-05
...

4.20 Bookend: Step 2

This is the end of this part.

Topic Selection

4.21 Score Computation (I)

Response Probabilities

We want a data file that looks like:

Item 1	Item 2	Item 3	...	Item 14
1	1	0	...	1
0	1	1	...	0
⋮	⋮	⋮	⋮	⋮
0	1	1	...	0

Let's consider one individual:

Theta	a_1	b_1	...	a_{14}	b_{14}
-1.43	1.7	-2	...	1.9	2

$$P_j(U_j = 1 | \omega_j, \theta) = \frac{1}{1 + e^{-a_j(\theta - b_j)}}$$

For item 1: $P(U_1 = 1) = \frac{1}{1 + e^{-1.7(-1.43 - (-2))}} = .725$

⋮

For item 14: $P(U_{14} = 1) = \frac{1}{1 + e^{-1.9(-1.43 - (-2))}} = .0015$



4.22 Score Computation (II)

Response Computation (I)

- Incorporate **randomness** into the process
- Choose a **random number** between 0 and 1 with **equal probability** from a **uniform distribution**
 - ✓ If the randomly selected value is **less than or equal to .725**, assign a score of **1**
 - ✓ If the randomly selected value is **greater than .725**, assign a score of **0**



For item 1: $P(U_1 = 1) = \frac{1}{1 + e^{-1.7(-1.43 - (-2))}} = .725$

4.23 Score Computation (III)

**Response Computation (II)**

```
data responses;
  set all_parms;
  array a{&n_item} a1-a&n_item;
  array b{&n_item} b1-b&n_item;
  array p{&n_item} p1-p&n_item;
  array item{&n_item} item1 - item&n_item;
```

4.24 Score Computation (IV)

**Response Computation (III)**

```
data responses;
  set all_parms;
  array a{&n_item} a1-a&n_item;
  array b{&n_item} b1-b&n_item;
  array p{&n_item} p1-p&n_item;
  array item{&n_item} item1 - item&n_item;
  do j = 1 to &n_item;
    p[j]=1/(1+exp(-(a[j]*(theta-b[j]))));
    y=rand('uniform');
    if y<=p[j] then item[j]=1;
    else if y>p[j] then item[j]=0;
  end;
run;
```

Results

Results (Slide Layer)

Results

Obs	p1	p2	p3	p4	p5	p6	p7	item1	item2	item3	item4	item5	item6	item7
20	0.96695	0.84241	0.69557	0.49407	0.29448	0.15139	0.03156	1	1	1	0	1	1	0
21	0.90752	0.64193	0.43382	0.24671	0.12279	0.05645	0.01081	0	1	0	0	0	0	0
22	0.74623	0.34946	0.18673	0.08937	0.04026	0.01761	0.00326	1	1	0	1	1	0	0
23	0.99891	0.99404	0.98616	0.96821	0.92866	0.84764	0.50406	1	1	1	1	1	1	1
24	0.98711	0.93327	0.85669	0.71870	0.52200	0.31822	0.07857	1	1	1	1	0	0	1
25	0.90371	0.63161	0.42290	0.23851	0.11807	0.05412	0.01034	1	1	0	1	0	0	0

Back

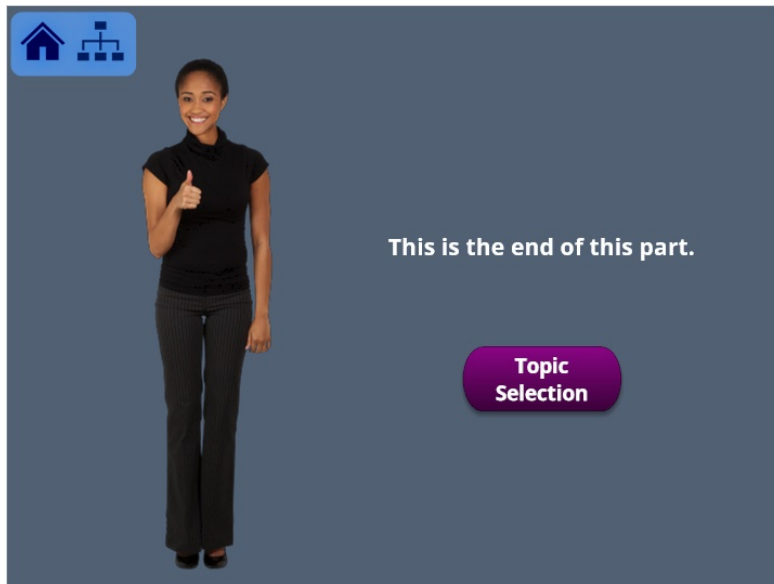
4.25 Score Computation (V)

Data Set Cleanup

```
data responses only;
set responses;
keep item:;
run;
```

	item1	item2	item3	item4	item5	item6	item7	item8	item9	item10
1	1	1	1	1	1	0	0	0	1	1
2	1	0	1	0	1	0	0	0	1	0
3	1	1	1	1	0	0	0	0	1	0
4	1	1	1	1	0	0	0	1	1	1
5	1	1	1	1	0	0	0	1	1	1
6	1	0	0	0	0	0	0	1	1	0
7	1	1	1	1	1	1	0	1	0	1
8	1	1	1	1	1	0	0	1	1	1
9	1	1	1	1	1	1	0	1	0	1
10	1	1	1	1	1	1	0	1	1	0
11	0	0	0	0	0	0	0	0	1	0
12	1	1	0	1	0	1	0	1	1	0
13	1	1	1	1	1	1	0	1	1	1
14	1	1	1	1	1	0	0	1	1	1
15	1	1	1	1	1	0	0	1	1	1
16	1	1	1	1	0	0	0	1	0	1

4.26 Bookend: Step 3



4.27 Parameter Validation (I)

Parameter Validation

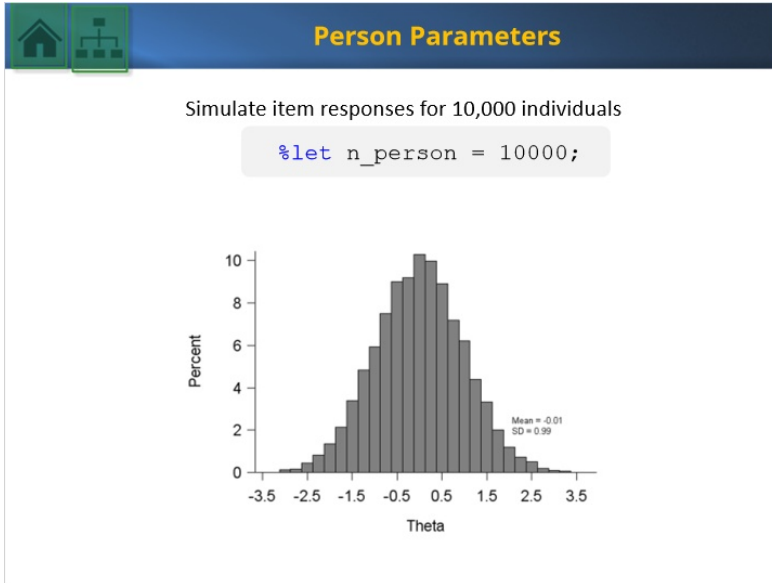
One common technique is to simulate scores for a large number of individuals (and number of items, if possible)

Two methods:

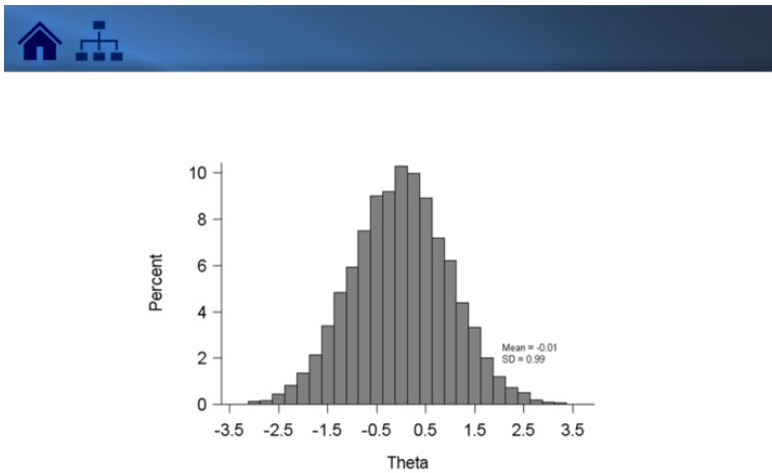
- Fit to the population
- Graphical

A large green checkmark is positioned to the right of the list of methods.

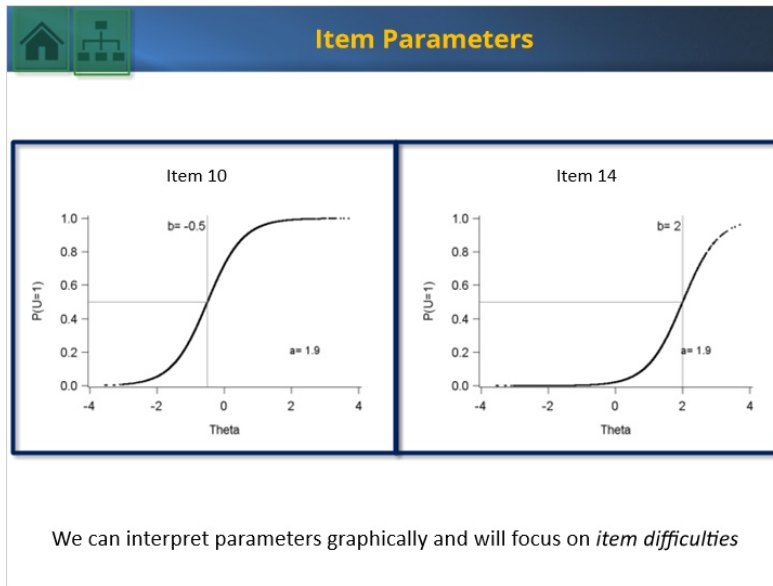
4.28 Parameter Validation (II)



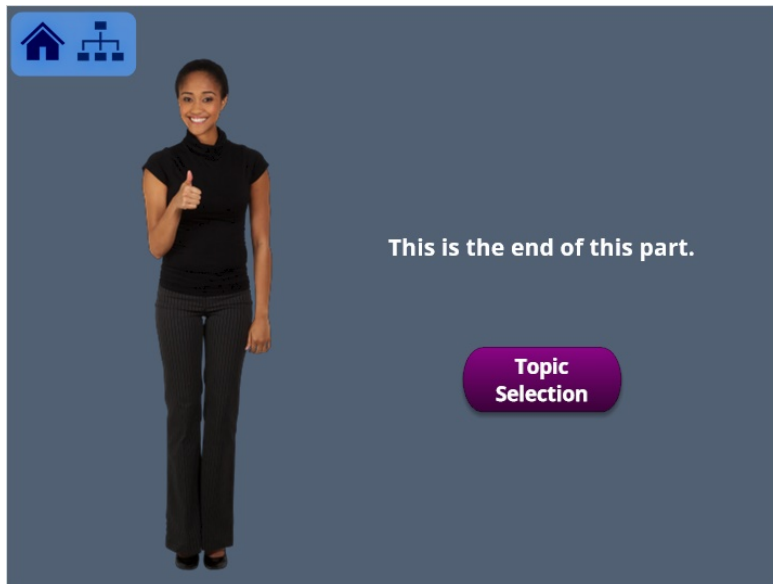
Results (Slide Layer)





4.29 Parameter Validation (III)



4.30 Bookend: Step 4




4.31 Replications (I)





Replications

“within each condition, there are often a certain number of replications. That is, for many research questions, it may be necessary to repeat the data generation process over multiple occasions—that is, replications—so that the empirical estimate of the sampling distribution (due to the simulation) of various statistics of interest may be observed”

Feinberg and Rubright, 2016

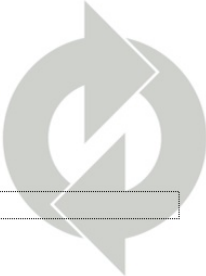


4.32 Replications (II)



Replications

```
%let n_reps=50;
%macro repeatIRTgen;
  %do rep = 1 %to &n_reps;
    [data generation commands]
  %end;
%mend repeatIRTgen;
%repeatIRTgen;
```



```
libname sim 'C:\Users\desktop\simulation';
data sim.Data&rep;
  set responses;
  keep item: theta;
run;
```

4.33 Replications (III)

What Do You Output?

What you save will depend on your research questions

Examples



- If you are looking at item parameter recovery
→ save true item parameters for each replication
- If you are investigating person parameter recovery
→ save true person ability parameters for each replication
- If you are investigating a new model that does not have the same parameters and are looking at a comparison between estimated expected score and true score
→ save the true scores of persons

4.34 Replications (IV)

General Setup

```
libname sim `C:/Users/[...]/Desktop/simulation';  
/*****macro variables *****/  
%let n_person = 100;  
%let n_reps=50;  
/*****  
data item_parms;  
  input a b;  
  cards;  
  1.7 -2  
  1.7 -1  
  1.7 -.5  
  1.7 0  
  1.7 .5  
  1.7 1  
  1.7 2  
  1.9 -2  
  1.9 -1  
  1.9 -.5  
  1.9 0  
  1.9 .5  
  1.9 1  
  1.9 2  
  ;  
run;  
proc sql noprint;  
select count(a) into :n_item trimmed from item_parms;  
run;
```

4.35 Replications (V)





Generating Person Parameters

```
%macro repeatIRTgen;
&do rep = 1 &to &n_reps;
proc datasets library=work noprint;
  save item_parms;
run;
quit;

data person_parm;
  do i = 1 to &n_person;
    theta=rand('normal',0,1);
    output;
  end;
  keep theta;
run;
```

Full code is available in the “Resources” of the module



4.36 Replications (VI)



Merging Item and Person Parameters

```
1  proc transpose data=item_parms out=a_wide prefix=a;
2    var a;
3  run;
4
5  proc transpose data=item_parms out=b_wide prefix=b;
6    var b;
7  run;
8
9  data items_wide;
10   merge a_wide b_wide;
11   drop _name_;
12 run;
13
14 data all_parms;
15   set person_parm;
16   if _N_=1 then set items_wide ;
17 run;
```



4.37 Replications (VII)



Simulating Responses

```
1 data responses;  
2   set all_parms;  
3   array a{&n_item} a1-a&n_item;  
4   array b{&n_item} b1-b&n_item;  
5   array p{&n_item} p1-p&n_item;  
6   array item{&n_item} item1 - item&n_item;  
7   do j = 1 to &n_item;  
8     p[j]=1/(1+exp(-(a[j]*(theta-b[j]))));  
9     y=rand('uniform');  
10    if y<=p[j] then item[j]=1;  
11    else if y>p[j] then item[j]=0;  
12  end;  
13 run;
```






4.38 Replications (VIII)



Saving Generated Data Sets

```
data sim.Data&rep;  
  set responses;  
  keep item: theta;  
run;  
  
%mend;  
%mend repeatIRTgen;  
  
%repeatIRTgen;
```

Name

-  data1.sas7bdat
-  data2.sas7bdat
-  data3.sas7bdat
-  data4.sas7bdat
-  data5.sas7bdat

4.39 Bookend: Step 5



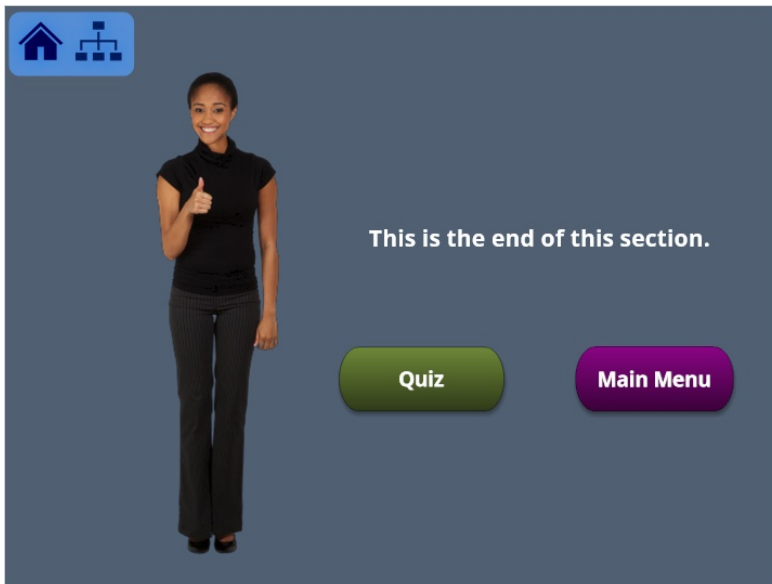
4.40 Other Useful Tips

Other Useful Tips

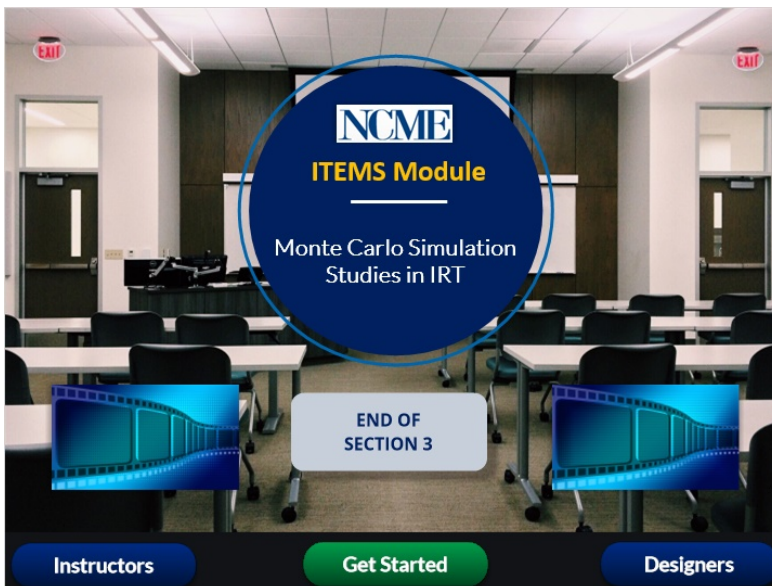
- You want your results to be reproduced
—> use a random seed
- Macros are very difficult to debug
—> develop macros only at the last step
- It is tedious to change after hard-coding values
—> start with macro variables
- Data validation usually doesn't go into a publication but is incredibly important
—> perform data validation
- Don't reinvent the wheel unnecessarily
—> call outside programs when appropriate

The slide has a dark blue header with the title 'Other Useful Tips' in yellow. Below the header are five light blue rounded rectangular boxes, each containing a bullet point and a recommendation. The boxes are arranged in a descending staircase pattern from top-left to bottom-right.

4.41 Bookend: Section 3



4.42 Module Cover (END)

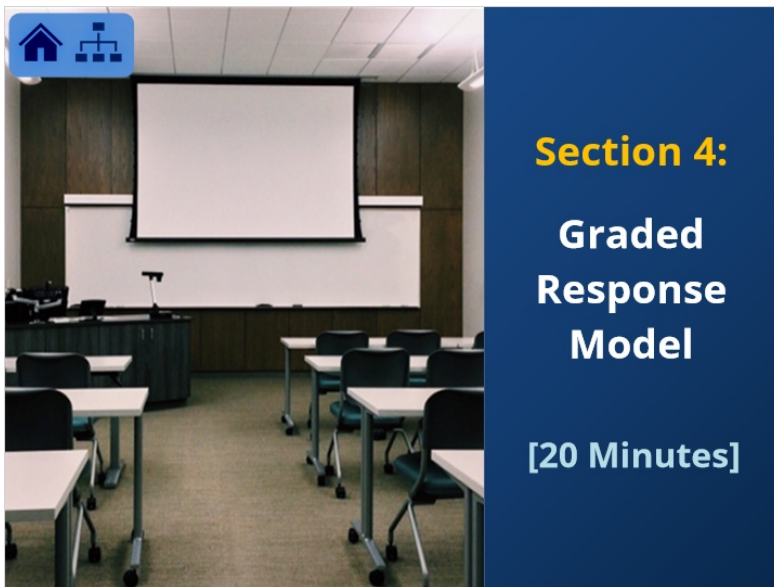


5. Section 4: Graded Response Model



5.1 Module Cover (START)




5.2 Cover: Section 4



5.3 Objectives





Learning Objectives



1. Identify key parameters in the graded response model
2. Describe how GRM simulation is an extension of 2PL simulation
3. Articulate the process for item response simulation using the GRM
4. Modify example code for specific research scenarios

5.4 Simulation Goal



Simulation Goal

- Simulate item responses for 500 individuals to a 10-item survey *according to the graded response model*
- Each item will be polytomously scored with responses 0, 1, 2, 3, and 4
- Ideally, we want to end with a data file that looks like:

Person	Item 1	Item 2	Item 3	...	Item 10
1	1	2	0	...	1
2	4	3	4	...	4
⋮	⋮	⋮	⋮	⋮	⋮
500	3	3	2	...	3

Each row represents an individual's response set



5.5 Topic Selection



5.6 Bookmark: GRM




5.7 GRM (I)

 Overview

Probabilities (p_{ijk}) for individual score categories are defined as differences between cumulative probabilities (p_{ijk}^*)



Example

$$p_{ijk=3}^* - p_{ijk=4}^* = p_{ijk=3}$$



Each p_{ijk}^* is modeled as a 2PL

5.8 GRM (II)

 Core Formula

$$P_j^*(U_j \geq k | \theta) = \frac{1}{1 + e^{-Da_j(\theta - b_{jk})}}$$

- Cumulative probability an individual with a given ability or trait θ selects or receives a category score of k or higher for item j
- U_j : response on item j
- θ is an ability value on the trait of interest
- a_j is the slope/discrimination parameter for item j
- b_{jk} is the threshold parameter for item j and category k
- D is the scaling parameter
 - ✓ Here we set $D = 1$
 - ✓ Other common value is $D = 1.702$

Reference

References (Slide Layer)



Handbook of Modern Item Response Theory pp 85-100 | [Cite as](#)

Graded Response Model

Authors [Authors and affiliations](#)

Fumiko Samejima

Chapter 302 1.7k
Citations Downloads

Abstract

The graded response model represents a family of mathematical models that deals with ordered polytomous categories. These ordered categories include rating such as letter grading, A, B, C, D, and F, used in the evaluation of students' performance; strongly disagree, disagree, agree, and strongly agree, used in attitude surveys; or partial credit given in accordance with an examinee's degree of attainment in solving a problem.

Keywords

[Latent Trait](#) [Item Parameter](#) [Homogeneous Case](#) [Partial Credit](#) [Grade Response Model](#)

These keywords were added by machine and not by the authors. This process is experimental and the keywords may be updated as the learning algorithm improves.

Back

5.9 GRM (III)



$$P_j^*(U_j \geq k | \theta) = \frac{1}{1 + e^{-Da_j(\theta - b_{jk})}}$$

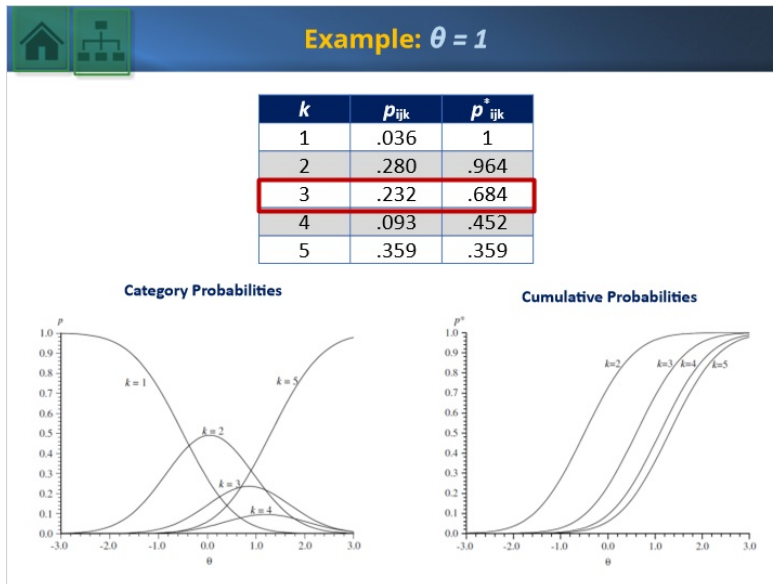
Item j with Categories $k = 0, 1, \dots, K$

$$P_j(U_j = 0) = 1 - P_j^*(U_j \geq 1 | \theta) \quad k = 0$$

$$P_j(U_j = k) = P_j^*(U_j \geq k | \theta) - P_j^*(U_j \geq k + 1 | \theta) \quad 1 < k < K - 1$$

$$P_j(U_j = K) = P_j^*(U_j \geq K | \theta) \quad k = K$$

5.10 GRM (IV)



5.11 GRM (V)

Thresholds

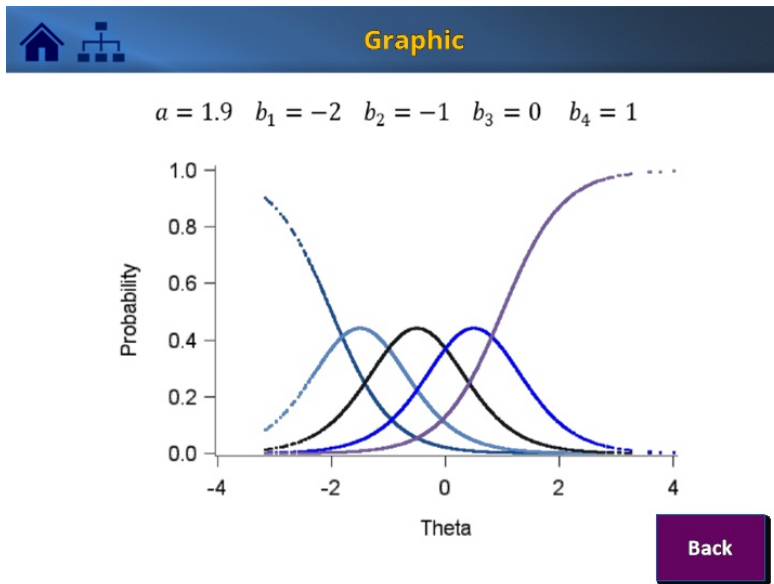
- b_{jk} is the threshold parameter for item j and category k
- b_{jk} is the point on the θ scale in which a respondent has a 0.5 probability of selecting that category or greater
- For an item with K categories, there are $K - 1$ thresholds

Example

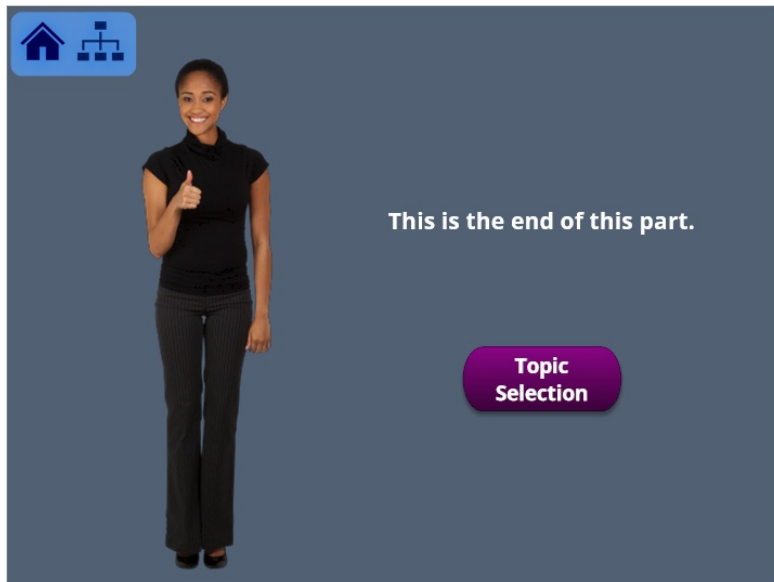
For an item with 5 categories scored 0, 1, 2, 3, and 4,
there are four thresholds such that:

$$b_{j1} \leq b_{j2} \leq b_{j3} \leq b_{j4}$$

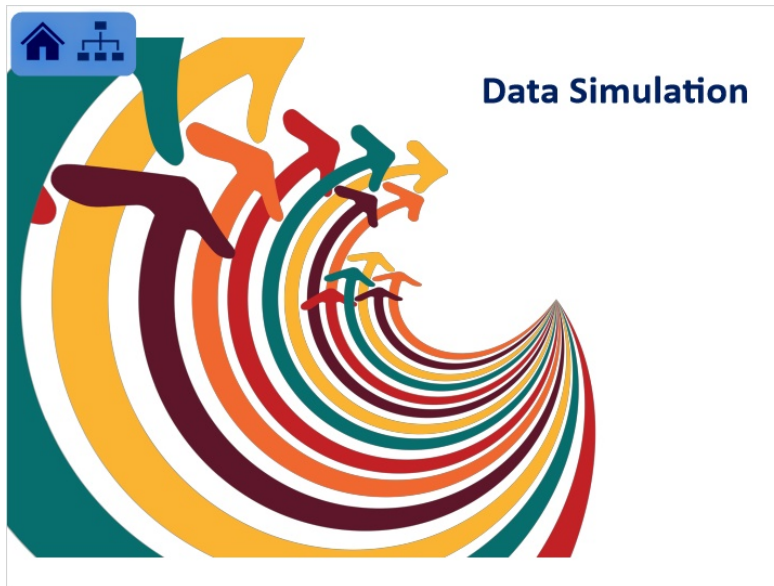
Graphic (Slide Layer)



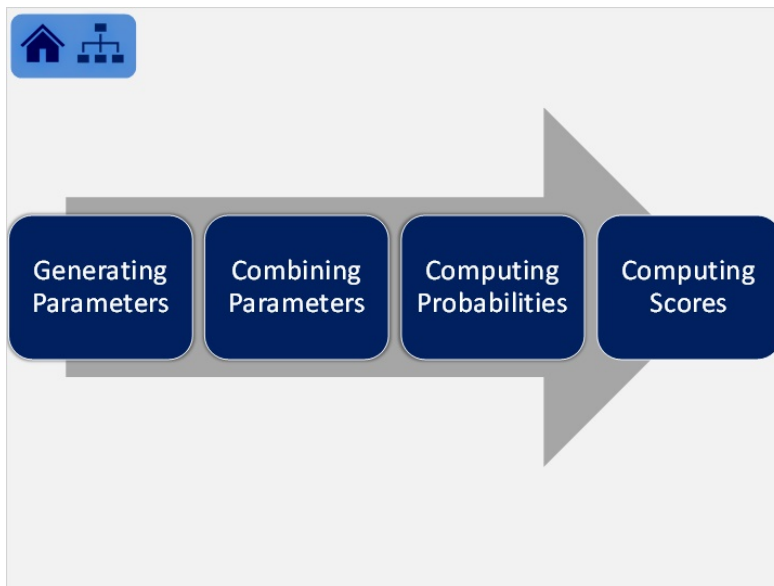
5.12 Bookend: GRM





5.13 Bookmark: Data Simulation



5.14 Step Selection



5.15 Person Parameters



 

Person Parameters

```
%let n_person =500;
data person_parm;
  do i = 1 to &n_person;
    theta=rand('normal',0,1);
    output;
  end;
  keep theta;
run;
```

Create a value named *theta*. Randomly select theta value from a normal distribution with mean = 0 and var = 1

5.16 Item Parameters

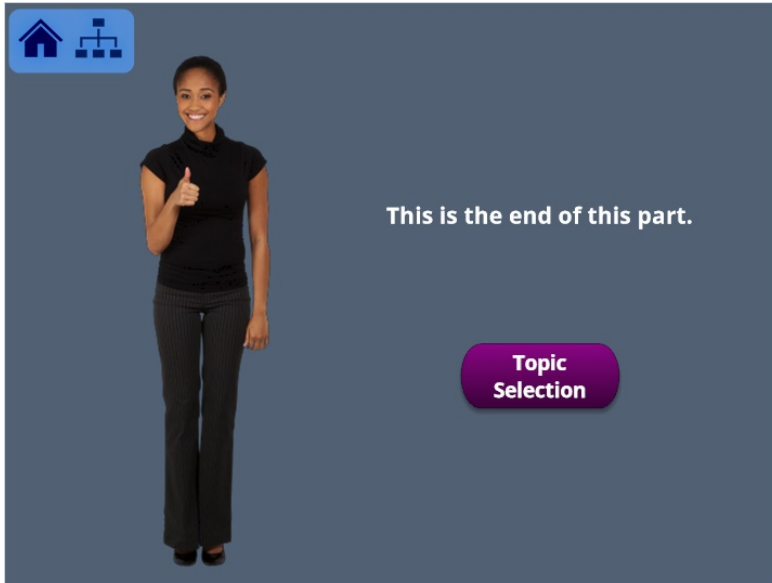
 

Item Parameters

- Create a data set that contains **item parameters**
- This may be based off of **empirical estimates, previous literature, etc.**
- You may **randomly select** item parameters as well
- **Literature** can be helpful in determining what **distribution** to randomly select from

```
data item_parms;
input a b1 b2 b3 b4;
cards;
1.2 -2 -1 0 1
1.2 -1.5 -.5 .5 1.5
1.2 -1 0 1 2
1.2 -.5 .5 1.5 2.5
1.2 0 .5 1 1.5
1.9 -2 -1 0 1
1.9 -1.5 -.5 .5 1.5
1.9 -1 0 1 2
1.9 -.5 .5 1.5 2.5
1.9 0 .5 1 1.5
run;
```

5.17 Bookend: Step 1



5.18 Parameter Matrix (I)

Parameter Matrix (I)

- We need in the same data set / parameter matrix:
 - ✓ an individual's ability parameter θ
 - ✓ item parameters a_j and b_{jk}
- We want our data set / parameter matrix to look like this:

Theta	a_1	$b_{1,1}$	$b_{1,2}$	$b_{1,3}$	$b_{1,4}$...	a_{10}	$b_{10,1}$	$b_{10,2}$	$b_{10,3}$	$b_{10,4}$
1.43	1.2	-2	-1	0	1	...	1.9	0	.5	1	1.5
2.81	1.2	-2	-1	0	1	...	1.9	0	.5	1	1.5
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.09	1.2	-2	-1	0	1	...	1.9	0	.5	1	1.5

5.19 Parameter Matrix (II)

Parameter Matrix (II)

```
proc transpose data=item_parms out=a_wide prefix=a;
var a;
run;

proc transpose data=item_parms out=b1_wide prefix=b1_;
var b1;
run;

proc transpose data=item_parms out=b2_wide prefix=b2_;
var b2;
run;

proc transpose data=item_parms out=b3_wide prefix=b3_;
var b3;
run;

proc transpose data=item_parms out=b4_wide prefix=b4_;
var b4;
run;

data items_wide;
merge a_wide b1_wide b2_wide b3_wide b4_wide;
drop name_;
run;
```

Turn
parameters
into wide form

5.20 Parameter Matrix (III)

Parameter Matrix (III)

```
data all_parms;
set person_parm;
if _N_=1 then set items_wide ;
run;
```

	ITEM	ITEM1	ITEM2	ITEM3	ITEM4	ITEM5	ITEM6	ITEM7	ITEM8	ITEM9	ITEM10	ITEM11	ITEM12	ITEM13	ITEM14	ITEM15	ITEM16	
1	0306191724	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
2	0306191727	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
3	0306191762	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
4	0306194023	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
5	0306194191	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
6	0306194197	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
7	0306194554	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
8	0306194565	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
9	0306194786	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
10	0306194821	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
11	0306194876	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
12	0306194882	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
13	0306194897	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
14	0306194946	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
15	0306194949	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
16	0306194995	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
17	0306195027	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
18	0306195023	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
19	0306195038	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
20	0306195049	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
21	0306195284	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
22	0306195285	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
23	0306194476	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
24	0306194746	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
25	0306194949	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
26	0306194949	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
27	0306194946	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15
28	0306194946	12	12	12	12	12	19	19	19	19	19	2	15	-1	05	0	2	-15

5.21 Bookend: Step 2



This is the end of this part.

Topic Selection

5.22 Cumulative Probabilities

Cumulative Probabilities


Theta	a_1	$b_{1,1}$	$b_{1,2}$	$b_{1,3}$	$b_{1,4}$...
1.43	1.2	-2	-1	0	1	...
2.81	1.2	-2	-1	0	1	...
...
0.09	1.2	-2	-1	0	1	...

$$P_j^*(U_j \geq k | \theta) = \frac{1}{1 + e^{-a_j(\theta - b_{jk})}}$$



$$P(U_1 \geq 1) = \frac{1}{1 + e^{-1.2(1.43 - (-2))}} = .984$$


$$P(U_1 \geq 2) = \frac{1}{1 + e^{-1.2(1.43 - (-1))}} = .949$$

$$P(U_1 \geq 3) = \frac{1}{1 + e^{-1.2(1.43 - (0))}} = .848$$



$$P(U_1 \geq 4) = \frac{1}{1 + e^{-1.2(1.43 - (1))}} = .626$$



5.23 Category Probabilities

  **Category Probabilities**

$$P_j(U_j = 0) = 1 - P_j^*(U_j \geq 1|\theta)$$
$$= 1 - .984 = .016$$
$$P_j(U_j = 1) = P_j^*(U_j \geq 1|\theta) - P_j^*(U_j \geq 2|\theta)$$
$$= .984 - .949 = .035$$
$$P_j(U_j = 2) = P_j^*(U_j \geq 2|\theta) - P_j^*(U_j \geq 3|\theta)$$
$$= .949 - .848 = .101$$
$$P_j(U_j = 3) = P_j^*(U_j \geq 3|\theta) - P_j^*(U_j \geq 4|\theta)$$
$$= .848 - .626 = .221$$
$$P_j(U_j = 4) = P_j^*(U_j \geq 4|\theta)$$
$$= .626$$


5.24 Bookend: Step 3



This is the end of this part.

Topic Selection

5.25 Response Computation (I)

Response Computation (I)

Prob of 0: .016
Prob of 1: .051
Prob of 2: .152
Prob of 3: .373
Prob of 4: .626

- Choose a **random number** between 0 and 1 with **equal probability** from a **uniform distribution**
- If the random number falls between response categories **k-1** and **k** then response category **k** is assigned to the item

$P_j(U_j = 0) = .016$
$P_j(U_j = 1) = .035$
$P_j(U_j = 2) = .101$
$P_j(U_j = 3) = .221$
$P_j(U_j = 4) = .626$

5.26 Response Computation (II)

Response Computation (II)


```
data responses;  
  set all_parms;  
  array a{&n_item} a1-a&n_item;  
  array b1_{&n_item} b1_1-b1_&n_item;  
  array b2_{&n_item} b2_1-b2_&n_item;  
  array b3_{&n_item} b3_1-b3_&n_item;  
  array b4_{&n_item} b4_1-b4_&n_item;  
  
  array p0{&n_item} p0_1-p0_&n_item;  
  array p1{&n_item} p1_1-p1_&n_item;  
  array p2{&n_item} p2_1-p2_&n_item;  
  array p3{&n_item} p3_1-p3_&n_item;  
  array p4{&n_item} p4_1-p4_&n_item;  
  
  array item{&n_item} item1 - item&n_item;
```

5.27 Response Computation (III)

Response Computation (III)

```

do j = 1 to &n item;
  p1_star=1/(1+exp(-(a[j]*(theta-b1_[j]))));
  p2_star=1/(1+exp(-(a[j]*(theta-b2_[j]))));
  p3_star=1/(1+exp(-(a[j]*(theta-b3_[j]))));
  p4_star=1/(1+exp(-(a[j]*(theta-b4_[j]))));

  p0[j]=1-p1_star;
  p1[j]=p1_star-p2_star;
  p2[j]=p2_star-p3_star;
  p3[j]=p3_star-p4_star;
  p4[j]=p4_star;

  score_temp=rand("Table", p0[j], p1[j], p2[j], p3[j], p4[j]);
  item[j]=score_temp-1;
end;

```

Calculate cumulative probabilities

Calculate score probabilities

Convert scores (0-4)

Assign scores (1-5)

5.28 Data Set Cleanup

Data Set Cleanup

Keep only variables with prefix "item"

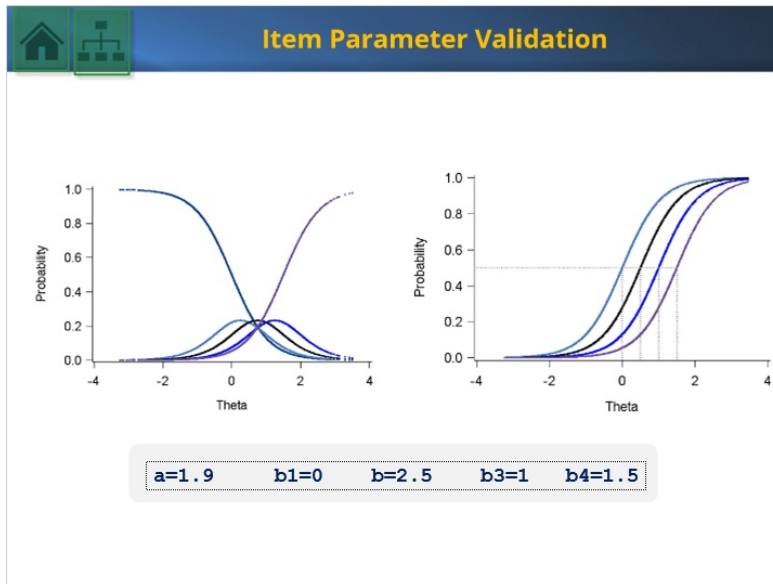
```

data responses_only;
set responses;
keep item:;
run;

```

	Item1	Item2	Item3	Item4	Item5	Item6	Item7	Item8	Item9	Item10	
1	4	3	3	4	4	4	4	4	3	2	3
2	4	3	2	2	4	4	4	4	1	3	3
3	4	3	2	2	4	4	3	4	3	4	4
4	3	2	2	0	0	3	1	1	0	0	0
5	4	3	3	1	0	3	3	2	3	0	0
6	3	3	2	1	0	4	3	2	1	4	4
7	3	1	2	1	1	2	3	3	2	1	1
8	3	2	0	0	4	2	3	0	0	0	0
9	3	1	1	3	0	1	0	1	2	0	0
10	1	2	0	1	4	3	1	2	2	0	0
11	2	4	4	4	4	4	4	1	4	1	1
12	3	2	2	2	0	3	3	3	1	3	3
13	1	3	0	0	2	0	0	1	0	0	0
14	2	4	0	1	1	1	2	3	1	1	1
15	3	0	0	0	0	3	0	0	0	0	0
16	1	0	0	0	0	3	0	0	0	0	0
17	3	1	3	0	0	2	2	0	2	0	0
18	2	4	1	0	3	1	3	4	1	2	2
19	1	2	0	0	1	0	0	0	0	0	0
20	0	3	0	0	0	0	0	0	0	0	0
21	2	2	1	3	3	2	3	2	3	3	3
22	2	1	1	0	0	3	2	0	1	4	4
23	2	3	2	0	1	1	2	0	1	4	4
24	4	4	3	1	0	3	4	2	2	3	3

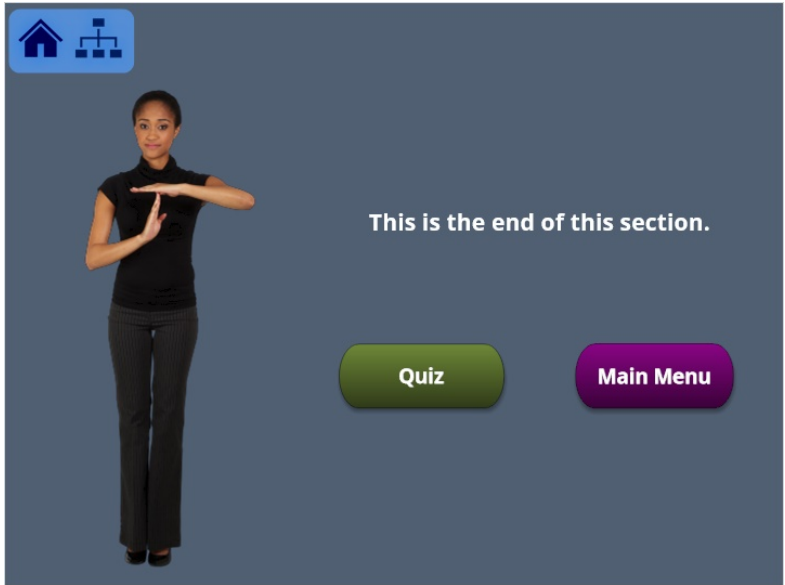
5.29 Item Parameter Validation



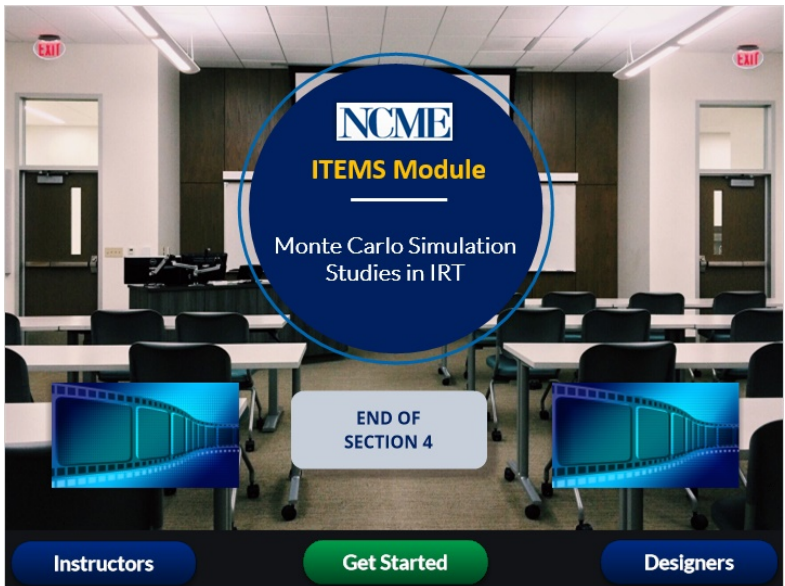
5.30 Bookend: Step 4

The slide titled "Bookend: Step 4" features a woman in a black top and pants giving a thumbs up. The text "This is the end of this part." is displayed to the right. A purple button labeled "Topic Selection" is located in the bottom right corner.

5.31 Bookend: Section 4



5.32 Module Cover (END)



6. Section 5: Bifactor Model [Advanced]



6.1 Module Cover (START)




6.2 Cover: Section 4



6.3 Objectives



Learning Objectives



1. Articulate research questions related to the bifactor model for simulation studies
2. Describe a data generation process for multidimensional data
3. Adapt sample code to conduct a simulation for bifactor models
4. Interpret results from a bifactor simulation

6.4 Topic Selection



Bifactor Model

Data Simulation

Section End

6.5 Bookmark: Bifactor Model

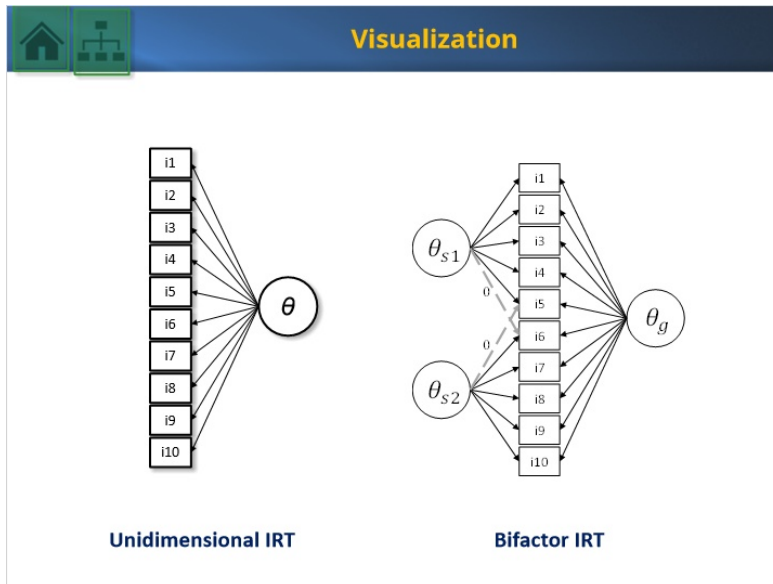


6.6 Bifactor Model (I)

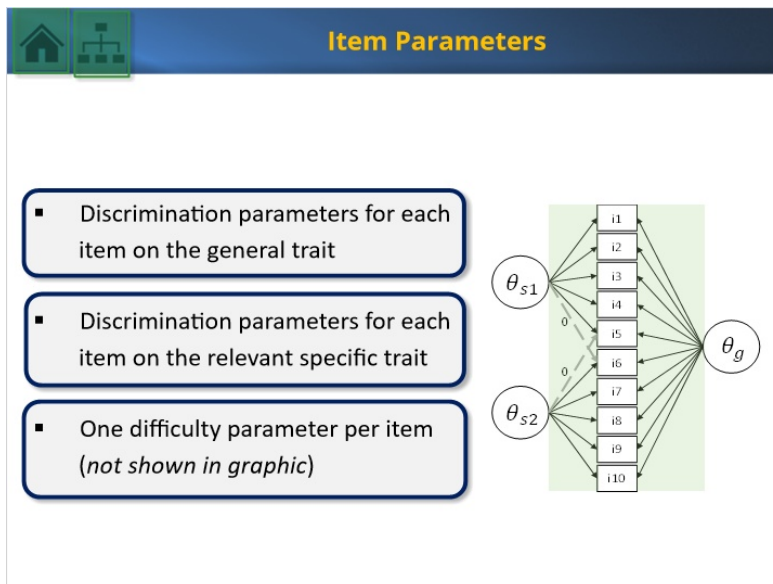
Overview

- Bifactor models are multidimensional
 - ✓ General factor: Primary trait of interest
 - ✓ Specific (secondary) factors: Capture the residual item relationships beyond those accounted by the general factor
- Each response is determined by the general and no more than one specific factor
- The specific factors are orthogonal to the general factor (uncorrelated)

6.7 Bifactor Model (II)



6.8 Bifactor Model (III)



6.9 Bifactor Model (IV)

Core Formula

$$P_j(U_j = 1 | \omega_j, \theta) = \frac{1}{1 + e^{-(a_{js1}\theta_{s1} + a_{js2}\theta_{s2} + a_{jg}\theta_g + d_j)}}$$

P_j : Response probability for item j

U_j : Response on item j
 1: correct
 0: incorrect

$U_j = 1$: We are specifically modeling the probability of a correct response

ω_j : Vector of all item parameters for item j

θ : Vector of person parameters (general and specific)

6.10 Bifactor Model (V)

Item Parameters

$$P_j(U_j = 1 | \omega_j, \theta) = \frac{1}{1 + e^{-(a_{js1}\theta_{s1} + a_{js2}\theta_{s2} + a_{jg}\theta_g + d_j)}}$$

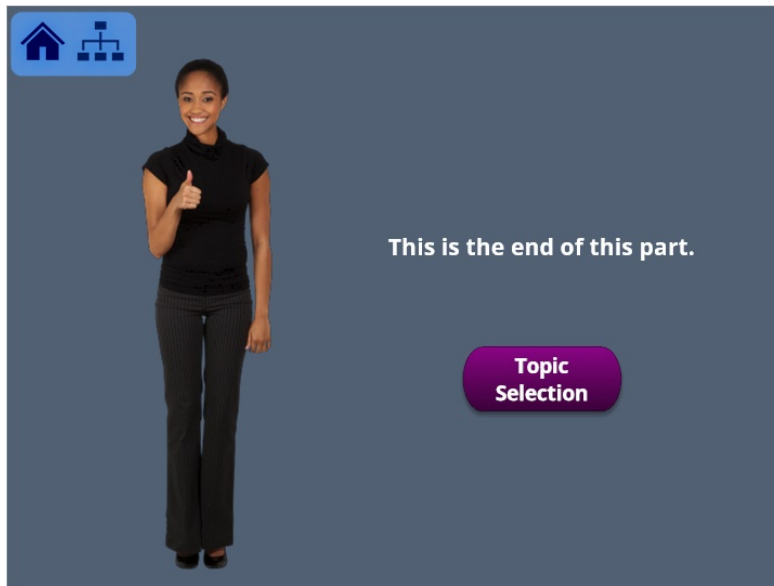
a_{js1} : discrimination parameter for item j on specific trait 1

a_{js2} : discrimination parameter for item j on specific trait 2

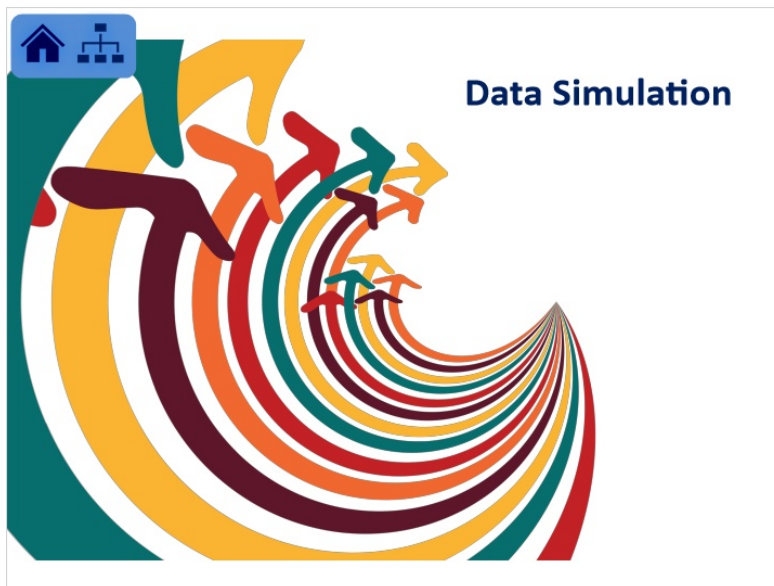
a_{jg} : discrimination parameter for item j on the general trait

d_j : difficulty / threshold parameter



6.11 Bookend: Bifactor Model



6.12 Bookmark: Data Simulation



6.13 Simulation Goal





Simulation Goal

- Simulate Item Responses for 1000 individuals to an 10-item test according to the *bifactor IRT model*
- Each item will be dichotomously scored as:
 - 1 = Correct
 - 0 = Incorrect
- Ideally, we want to end with a data file that looks like:



Each row represents an individual's response set

Item 1	Item 2	Item 3	...	Item 10
1	1	0	...	1
0	1	1	...	0
⋮	⋮	⋮	⋮	⋮
0	1	1	...	0

6.14 Step Selection





Simulation Steps

1. Specifying the IRT research question(s) 
2. Defining and justifying conditions
3. Specifying the experimental design and outcome(s) of interest
4. Simulating data under the specified conditions 
5. Estimating parameters
6. Comparing true and estimated parameters
7. Replicating the procedure a specified number of times
8. Analyzing results based on the design and research questions

[Click on each row to learn more](#) Topic Selection

6.15 Research Questions



Step 1: Research Question(s)

- **Example:** [Cai, L., Yang, J., & Hansen, M. \(2011\). Generalized full-information item bifactor analysis. *Psychological Methods*, 16\(3\), 221-248.](#)
Research Question: What are the bias and precision of the item parameters from a mixed format (dichotomous and polytomously scored) bifactor assessment?
Parameter Matrix (I)
- **Example:** [Fukuhara, H., & Kamata, A. \(2011\). A bifactor multidimensional item response theory model for differential item functioning analysis on testlet-based items. *Applied Psychological Measurement*, 35\(8\), 604-622.](#)
Research Question: What is the accuracy of estimating DIF magnitude by the proposed bifactor MIRT DIF model?



6.16 Bookend: Step 1



This is the end of this step.

Step Selection


6.17 Justifying Conditions





Step 2: Justifying Conditions

- Cai, Yang, & Hansen (2011)

Discrimination on the dichotomous and polytomously scored items
- Our simple example will manipulate sample size
 $n = 500$
 $n = 2000$




6.18 Bookend: Step 2



This is the end of this step.

Step Selection

6.19 Experimental Design





Step 3: Experimental Design

- Cai, Yang, & Hansen (2011):
Bias and estimated standard errors of the item parameters
- Our simple example will focus on **bias** of the **item parameters**

$$bias = \frac{\sum_{i=1}^n (\hat{\omega}_i - \omega_{true})}{n}$$

- $\hat{\omega}_i$ is the estimated item parameter for replication i
- ω_{true} is the generating item parameter
- n is the number of replications

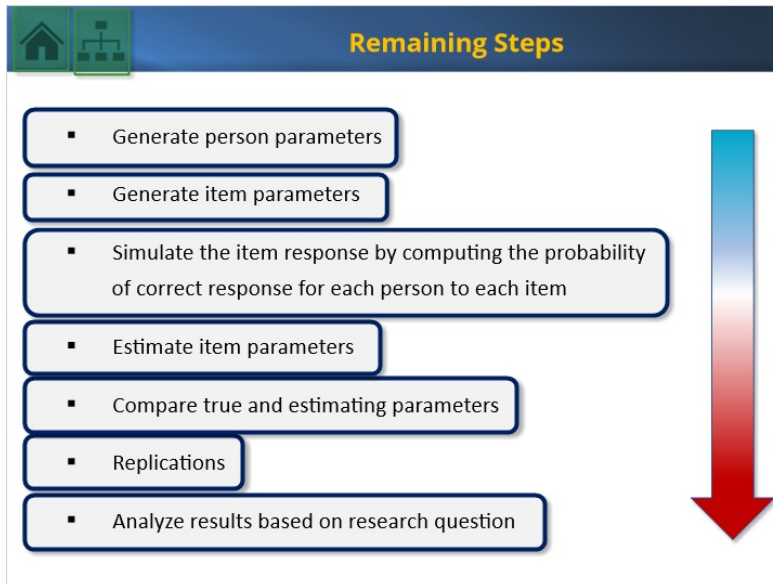
6.20 Bookend: Step 3



This is the end of this step.

Step Selection

6.21 Remaining Steps



6.22 Person Parameters

The slide titled "Step 4: Person Parameters" features a dark blue header with a home icon and a tree icon. The SAS code is displayed in a light gray rounded rectangle.

```
%let n_person = 500;
data person_parm;
  do i = 1 to &n_person;
    thetas=rand('normal',0,1);
    thetam=rand('normal',0,1);
    thetag=rand('normal',0,1);
    output;
  end;
  keep thetas thetam thetag;
run;
```

6.23 Item Parameters

Step 4: Item Parameters

- Borrow item parameters from DeMars (2013)
- Loadings not in the model are fixed to 0

Item	a_{i1}	a_{i2}	a_{i3}	d_i
1	1.06	0.44	0	-0.01
2	1.59	0.45	0	-0.13
3	0.98	0.36	0	0.44
4	0.88	0.48	0	-0.34
5	0.27	0.33	0	1.26
6	1.04	0	0.51	-0.77
7	0.61	0	0.23	-0.07
8	1.34	0	0.45	-0.14
9	0.96	0	0.46	2.22
10	1.20	0	0.42	0.18



6.24 Parameter Matrix (I)

Step 4: Parameter Matrix (I)

```

data item_parms;
input ag as am d;
cards;
1.06 .44 0 -.01
1.59 .45 0 -.13
0.98 .36 0 0.44
0.88 .48 0 -.34
0.27 .33 0 1.26
1.04 0 .51 -.77
0.61 0 .23 -.07
1.34 0 .45 -.14
0.96 0 .46 2.22
1.20 0 .42 0.18
;
run;
```

6.25 Parameter Matrix (II)

 **Step 4: Parameter Matrix (II)**



▪ We want our data set / parameter matrix to look like this:

Thetag	Thetas	Thetam	a_g1	a_s1	a_m1	d_1	...
-1.43	1.02	0.22	1.06	0.44	0	-2	...
2.81	-0.05	0.57	1.06	0.44	0	-2	...
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0.09	1.3	-0.25	1.06	0.44	0	-2	...

▪ Each row represents an individual

▪ Individuals are all answering the same items so the item parameters are the same for each line

6.26 Parameter Matrix (III)

 **Step 4: Parameter Matrix (III)**

```
proc transpose data=item_parms out=as_wide prefix=as;
var as;
run;

proc transpose data=item_parms out=am_wide prefix=am;
var am;
run;

proc transpose data=item_parms out=ag_wide prefix=ag;
var ag;
run;

proc transpose data=item_parms out=d_wide prefix=d;
var d;
run;

data items_wide;
merge as_wide am_wide ag_wide d_wide;
drop _name_;
run;
```

6.27 Parameter Matrix (IV)

Step 4: Parameter Matrix (IV)

Let's consider one individual and the first item:

thetas	thetam	thetag	as1	am1	ag1	d1
1.54	-1.30	-0.81	0.44	0	1.06	-0.01

We want a data file that looks like this:

Item 1	Item 2	Item 3	...	Item 10
1	1	0	...	1
0	1	1	...	0
⋮	⋮	⋮	⋮	⋮
0	1	1	...	0

6.28 Parameter Matrix (V)

Step 4: Parameter Matrix (V)

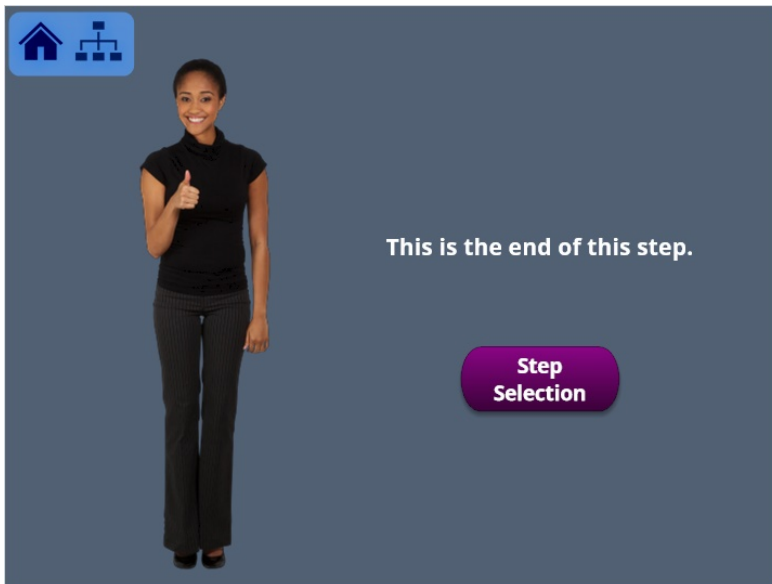
```

data responses;
set all_parms;
  array as{&n_item} as1-as&n_item;
  array am{&n_item} am1-am&n_item;
  array ag{&n_item} ag1-ag&n_item;
  array d{&n_item} d1-d&n_item;
  array p{&n_item} p1-p&n_item;
  array item{&n_item} item1 - item&n_item;
do j = 1 to &n_item;
  p[j]=1/(1+exp(-(as[j]*thetas+am[j]*thetam+ag[j]*thetag+d[j])))
  y=rand('uniform');
  if y<=p[j] then item[j]=1;
  else if y>p[j] then item[j]=0;
end;
run;

```

$$P_i(U_i = 1 | \omega, \theta) = \frac{1}{1 + e^{-(a_{s1}\theta_s + a_{m1}\theta_m + a_{g1}\theta_g + d_1)}}$$

6.29 Bookend: Step 4



6.30 Parameter Estimation

Step 5: Parameter Estimation

Estimate the bifactor model in *PROC IRT*

```
proc irt data=responses;
  var item1-item10;
  factor  Factorg===> item1-item10,
          Factors===> item1-item5= 1. load2 load3
                      load4 load5,
          Factorf===> item6-item10= 1. load7 load8
                      load9 load10;
  ods output ParameterEstimates=d1 slope=a1;
run;
quit;
```



6.31 Bookend: Step 5



6.32 Parameter Comparison (I)

```
data bias;  
merge item_parms d1 a1;  
biasd=d-Estimate;  
biasag=ag-Factorg;  
biasas=as-Factors;  
biasam=am-Factorm;  
item=_n_;  
run;
```

6.33 Parameter Comparison (II)



  **Step 6: Parameter Comparison (II)**

Items	ag	biasag	as	biasas	am	biasam	d	biasd
1	1.06	-0.1135	0.44	-0.56	-	-	-0.01	0.00695
2	1.59	-0.0178	0.45	0.16476	-	-	-0.13	-0.0107
3	0.98	0.03178	0.36	0.11529	-	-	0.44	0.01344
4	0.88	-0.0007	0.48	0.12659	-	-	-0.34	0.0003
5	0.27	-0.0299	0.33	0.18134	-	-	1.26	0.00216
6	1.04	-0.1116	-	-	0.51	-0.49	-0.77	0.11083
7	0.61	-0.0184	-	-	0.23	0.12497	-0.07	-0.0003
8	1.34	-0.0416	-	-	0.45	0.22995	-0.14	0.00957
9	0.96	-0.0267	-	-	0.46	0.23003	2.22	0.026
10	1.2	0.01481	-	-	0.42	0.14133	0.18	-0.0067


6.34 Bookend: Step 6





6.35 Replications (I)

**Step 7: Replications**


```
%macro IRTsim;  
%do rep = 1 %to &n_reps;  
  [Data generation steps]  
  ...  
  [Model estimation steps]  
  [Compute bias steps]  
%end;  
%mend IRTsim;  
%IRTsim;
```





6.36 Replications (II)

**Step 7: Keeping Track of Conditions (I)**

```
%macro IRTsim;  
%do np = 500 %to 2000 %by 1500;  
%do rep = 1 %to &n_reps;  
  [Data generation steps]  
  ...  
  [Model estimation steps]  
  [Compute bias steps]  
%end;  
%end;  
%mend IRTsim;  
%IRTsim;
```




6.37 Replications (III)

**Step 7: Keeping Track of Conditions (II)**

```
data responses; set all_parms;
  array as{&n_item} as1-as&n_item;
  array am{&n_item} am1-am&n_item;
  array ag{&n_item} ag1-ag&n_item;
  array d{&n_item} d1-d&n_item;
  array p{&n_item} p1-p&n_item;
  array item{&n_item} item1 - item&n_item;
  do j = 1 to &n_item;
    p[j]=1/(1+exp(-(as[j]*thetas+am[j]*thetam+ag[j]*thetad[j])));
    y=rand('uniform');
    if y<=p[j] then item[j]=1;
    else if y>p[j] then item[j]=0;
  end;
  drop j;
  rep=&rep;
  n=&np;
run;
```

6.38 Bookend: Step 7

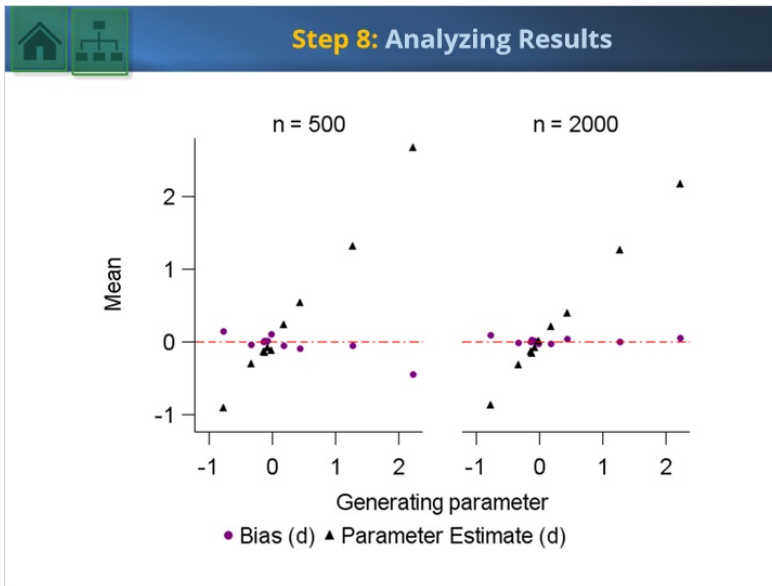




This is the end of this step.

Step Selection

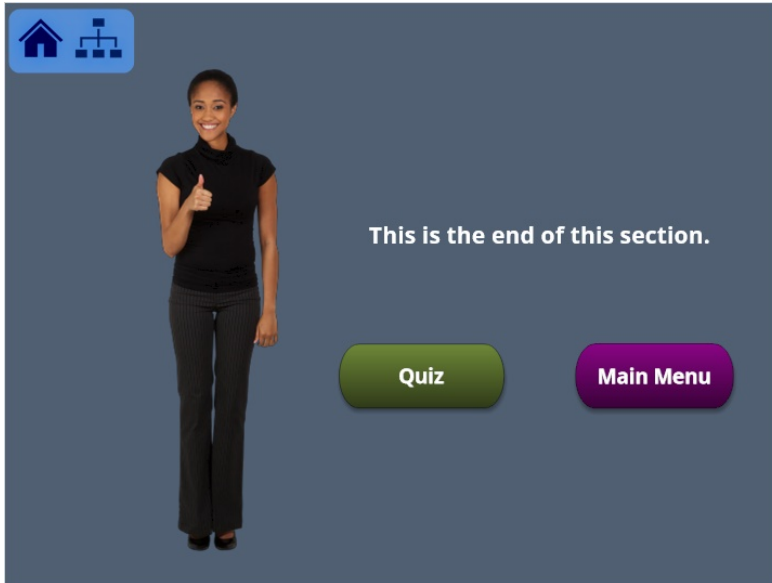
6.39 Analyzing Results



6.40 Bookend: Step 8



6.41 Bookend: Section 5



6.42 Module Cover (END)

